9

FG.

SCHOOL
OF
OPERATIONS RESEARCH
AND
INDUSTRIAL ENGINEERING

COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK

(1) TECHNICAL REPORT NO. 310

(11) Sep[redacted] 1976

(12) 107p.

(14) TR-310

GOODNESS OF FIT TESTS FOR COMPOSITE HYPOTHESES BASED ON AN INCREASING NUMBER OF ORDER STATISTICS

by

(10) Ray Hamilton Jakobovits

(15) Prepared under contracts

DAHC04-73-C-0008, U.S. Army Research Office - Durham,

N00014-75-C-0586; Office of Naval Research, and

under NSF GRANT NO. MCS76-06340.

THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS.

TABLE OF CONTENTS

CHAPTER I

Introduction

1.1 Introduction and Summary

A major area of continuing statistical research has been the development of procedures to test the conformity of data to a given hypothesis. Such procedures are known as goodness of fit tests. The hypothesis of interest usually stems from the assumption that the sample population is characterized by the same properties which characterize a particular distribution (simple hypothesis) or a particular family of distributions (composite hypothesis). Goodness of fit tests are useful in demonstrating by means of formal significance statements that such hypotheses are inadequate to explain the observations, or, in the absence of such significance statements, providing the impetus for further data analysis based on other, parametric, statistical techniques.

In this thesis, we are concerned with tests of the composite null hypothesis that the distribution function of the observed sample is a member of a specified scale-location parameter family, where the scale and location parameters are unknown and unspecified. We first study a test of this hypothesis proposed by Weiss [41]. Then we turn our attention to the most common special case--that of normality--by proposing and analyzing several tests which are large-sample analogs of the most efficacious (based on empirical evidence) small-sample procedures. All of our results are asymptotic, as the sample size n tends to infinity, and all of the test criteria which we examine are based on functions of an increasing subset of the ordered sample values.

By increasing subset, we mean that as the sample size tends to infinity, the number of order statistics which we consider also tends to infinity. By assuming certain relationships between the number of sample quantiles in the selected subset, the separation between those quantiles, the rate at which the quantiles move into the tails of the distribution, and the smoothness and size of the tails of the density function of a standard representative of the hypothesized scale-location parameter family, we are able to assert that the selected subset of sample quantiles is asymptotically jointly normally distributed with known mean vector and covariance matrix. The motivation behind choosing a gradually increasing subset is that we achieve a workable distribution theory while retaining "most" of the information contained in the sample.

The Weiss procedure is based on a quadratic form of the sample quantiles which has an asymptotic chi-square distribution. Weiss proved that, under the null hypothesis, replacing the unknown scale and location parameters by estimates obtained from the sample does not alter the asymptotic distribution theory, and he also proved that his procedure is consistent. In Chapter II, we examine the asymptotic power of the test under sequences of contiguous alternatives of the form $H_n(x) = G(\frac{x - \overline{\theta}_1}{\theta_2}) + \theta_n(x)$, where $G(y)$ is the standard representative of the scale-location parameter family appearing in the null hypothesis and the "disturbance function" $\varepsilon_n(x)$ satisfies certain regularity conditions. By letting the sequence $\{\varepsilon_n(x)\}$ approach zero at a certain rate and with respect to a certain measure of distance, we obtain a non-trivial power for the test.

In section 2.1, we explicitly define what we choose as the selected subset of order statistics, including those assumptions required for the asymptotic distribution theory. We also discuss the concept of asymptotic

indistinguishability of sequences of distribution functions which provides the basis for that theory. In section 2, we describe the Weiss test of fit. In section 2.3, we introduce the sequence $\{H_n(x)\}$, state the assumptions which the functions $\{\varepsilon_n(x)\}$ must satisfy to yield a non-trivial power, and derive several useful results based on those assumptions, including the joint asymptotic normality of the selected order statistics under the sequence of alternatives. Section 2.4 contains the derivation of the asymptotic power of the test. Finally, in section 2.5 we apply the results of section 2.4 to two specific cases--a sequence of Weibull distributions which approaches the exponential distribution and a sequence of contaminated normal distributions.

In Chapter III, we turn to the composite hypothesis of normality. Using the same approach of selecting a gradually increasing subset of order statistics, we develop large-sample analogs of the Wilk-Shapiro and Shapiro-Francia tests, two of the most effective (small samples) but also least understood tests for normality. We show that the analog of the Wilk-Shapiro statistic, which we call $\tilde{W}(n)$, is asymptotically normally distributed under the null hypothesis as the sample size tends to infinity. We also show that, asymptotically, the analog of the Shapiro-Francia statistic, which we call $\tilde{W}^*(n)$, suitably standardized, has the same distribution under H_0 as a certain weighted sum of independent chi-square random variables. We further prove that the tests based on $\tilde{W}(n)$ and $\tilde{W}^*(n)$ are consistent. Finally, we study the behavior of those tests under the sequence of alternatives $\{H_n(x)\}$ of Chapter II, except that here the sequences of disturbance functions must satisfy different assumptions regarding the rate and manner of approach to the null hypothesis in order to guarantee non-trivial powers. We find that the measures of distance,

or metrics, for the tests based on $\tilde{W}(n)$ and $\tilde{W}^*(n)$ are quite complicated. As a consequence of the power studies, we propose and analyze two additional statistics, $\tilde{\tilde{W}}(n)$ and $\tilde{\tilde{W}}^*(n)$, which yield improved tests for normality. We also find that, contrary to most small-sample empirical studies, the analogs of the Shapiro-Francia test have better asymptotic power than the analogs of the Wilk-Shapiro test, and we show by example that the analogs of the Wilk-Shapiro test can be biased, at least if the selected subset of order statistics is not chosen carefully.

Summarizing Chapter III, section 3.1 contains a survey of the literature regarding tests for normality. In section 3.2, we define the statistic $\tilde{W}(n)$ and determine its asymptotic distribution under the null hypothesis. We also state in section 3.2.4 the null hypothesis result for $\tilde{W}^*(n)$, deferring the proof to Appendix B since it depends on several results which are not used elsewhere in the thesis. In section 3.3, we prove that the $\tilde{W}$-test is consistent. In section 3.4, we analyze the power of the $\tilde{W}$-test under the sequence of alternatives $\{H_n(x)\}$. The analogous results for $\tilde{W}^*(n)$, and are given in section 3.5. In section 3.6, we illustrate the foregoing results using the example of a sequence of contaminated normal distributions and also compare the tests for normality with the Weiss test discussed in Chapter II. Example 3.2 illustrates the bias of $\tilde{W}(n)$ and $\tilde{\tilde{W}}(n)$. Finally, in section 3.7, we discuss possible areas of future research.

In Appendix A, we present the results of a Monte Carlo study of the Weiss statistic under the null hypothesis. In Appendix B, we present a linear transformation of the selected subset of order statistics to standard normal random variables and derive the asymptotic distribution of $\tilde{W}^*(n)$ under the null hypothesis as mentioned above. In Appendix C we

summarize for reference purposes various formulas concerning the inverse normal distribution function which are used frequently in Chapters II and III.

The remainder of this introduction is devoted to a brief discussion of some of the commonly used goodness of fit tests for continuous distributions, with emphasis on tests for the null hypothesis that the sample distribution is an unspecified member of a specified scale-location parameter family.

1.2 EDF Tests of Fit

A frequently used class of tests for simple or composite hypotheses about continuous distributions is the class of EDF tests, so-named because they are based on differences between the hypothesized distribution $G(x)$ and the empirical distribution function $G_n(x)$. Included in this class are tests based on the following statistics:

Kolmogorov-Smirnov: $D = \sup_x |G_n(x) - G(x)|$

Cramér-von Mises: $W^2 = n\int_{-\infty}^{\infty}[G_n(x) - G(x)]^2 dG(x)$

Anderson-Darling: $A^2 = n\int_{-\infty}^{\infty}[G_n(x) - G(x)]^2[G(x)(1 - G(x))]^{-1} dG(x)$

Letting $F(x)$ denote the unknown distribution function of the sample population, research on the above tests has been primarily devoted to the following cases, classified by the null hypotheses about $F(x)$: ($G(x)$ is continuous in all cases)

Case 0: $F(x) = G(x)$ completely specified

Case 1: $F(x) = G(x;\theta)$ where θ is a scalar-valued nuisance parameter

- Case 1A: $G(.)$ is normal with unknown mean
- Case 1B: $G(.)$ is normal with unknown variance
- Case 1C: $G(.)$ is exponential with unknown scale parameter

<u>Case 2</u>: $F(x) = G((x-\theta_1)/\theta_2)$ where θ_1 and θ_2 are nuisance location and scale parameters

<u>Case 2A</u>: $G(.)$ is normal with unknown mean and variance

<u>Case 2B</u>: $G(.)$ is exponential with unknown scale and location parameters (Durbin [12] shows how this case can be transformed to case 1C above.)

Case 0 is the classical case and has been extensively studied. The sampling distributions and asymptotic distributions of the statistics D, W^2, and A^2 under the null hypothesis are known to be independent of the hypothesized $G(x)$. A summary of the r.. lts and references appear in Darling [7]. Expressions for the powers of .e above EDF tests do not exist in closed form. Durbin [11] has shown how the power may be approximated in the case of D by computing boundary crossing probabilities for the tied-down Wiener process, and he has examined the asymptotic power under contiguous alternatives with densities of the form $h_n(x) = 1 + rn^{-1/2}$ when $G(x)$ is the uniform distribution. Empirical results for the powers of the various tests are scattered throughout the literature.

When nuisance parameters are present, the EDF tests are no onger distribution-free. The test based on D has been studied under the null hypothesis for case 2 by Serfling and Wood [24] and cases 1 and 2 (and the more general p-nuisance parameter case as well) by Durbin [12]. Serfling and Wood prove the weak convergence of the distribution function of the statistic to that of a certain Gaussian process with a covariance function which depends on $G(x)$ and the parameters being estimated. Durbin has developed techniques for inverting Fourier transforms to obtain the distribution of D for finite n and has worked out case 1C in detail.

Using Monte Carlo methods, Lilliefors [16,17] constructed tables of critical points for the Kolmogorov-Smirnov test in cases 1C and 2A. He found that the statistic D tended to be stochastically smaller when the unknown parameters were estimated from the sample.

The statistic W^2 has been studied under the null hypothesis in case 1 by Darling [6], in case 2A by Kac, Kiefer, and Wolfowitz [15], and both W^2 and A^2 have been studied by Stephens [34] for cases 1A - 1C and 2A. Briefly, the asymptotic distribution of each of the statistics under the null hypothesis is the same as that of the random variable

$$\int_0^1 Y^2(t)dt$$

where $Y(t)$ is an appropriate tied-down Gaussian process with zero mean and covariance function depending on the statistic, the hypothesized distribution and the parameters being estimated. This distribution is the same as that of

$$\sum_1^\infty Z_i/\lambda_i$$

where Z_i are i.i.d. chi-square random variables with 1 degree of freedom and λ_i are the eigenvalues of a certain integral equation. In [30], Stephens gives modifications of the statistics D, W^2, and A^2 for cases 0, 1A-1C, and 2A so that the asymptotic critical points of the associated tests may be used for finite sample sizes.

Kac, et. al. studied the asymptotic power of tests based on D and W^2 by examining the ways in which those tests measure distance, showing, in particular, that those tests are hyper-efficient relative to the optimum

chi-square test. Stephens [30], in empirical power studies on tests based on D, W^2, A^2, and two other EDF statistics, found that it was better <u>not</u> to have the true parameter values available in conducting the tests. Apparently, in trying to fit a density of a certain shape to the data, being tied down to fixed parameter values, even correct ones, is more a hindrance than a help.

1.3 Other Tests

The oldest well-established test of fit is the chi-square test. It is particularly well suited to the case where $F(x)$ is discrete, and it is known how to adapt the statistic to the case when there are nuisance parameters which must be estimated from the sample. When $F(x)$ is continuous, however, chi-square tests are neither consistent nor do they perform as well in general as the EDF tests mentioned above. A unified summary of chi-square theory appears in the paper by Moore and Spruill [19].

Sample spacings, or differences between successive ordered sample values, have provided the basis for many tests of fit. A comprehensive review of sample spacings theory, associated tests of fit, and the related literature is given by Pyke [20]. Under the null hypothesis that $F(x)$ is uniform, Cibisov [4] proved that the Kolmogorov-Smirnov test is hyper-efficient relative to any test symmetric in the sample spacings under sequences of continuous, contiguous alternatives of the form $h_n(x) = 1 + r(x)n^{-1/2}$. However, in his review of Cibisov's paper, Weiss [37] exhibited a sequence of alternatives such that the reverse is true. Sethuraman and Rao [25] studied the Pitman (asymptotic relative) efficiencies of various tests based on sample spacings for alternatives

$h_n(x) = 1 + r_n(x)n^{-\delta}$ for $\delta \geq 1/4$. And Blumenthal [2] studied the large-sample behavior of tests based on spacings when nuisance location and scale parameters are present.

Finally, there are many tests of fit which have been developed to test a specific null hypothesis, the most common being that of normality. A detailed survey of tests for normality is presented in section 3.1. Most of these tests share the property of location and scale invariance, thus eliminating the need for estimators of the nuisance parameters when the null hypothesis is composite.

CHAPTER II

ON THE WEISS TEST OF FIT

2.0 Introduction

For each n, $X_1(n), X_2(n), \ldots, X_n(n)$ are independent, identically distributed random variables with distribution function $F(x)$. $F(x)$ is known to be absolutely continuous with density function denoted by $f(x)$ but is otherwise unknown. Formally, the decision problem of interest is the following:

Problem P_n: Using $X_1(n), X_2(n), \ldots, X_n(n)$, test the hypothesis

$$H_0\colon \ F(x) = G((x-\theta_1)/\theta_2) \ \text{ for some } \ \theta_1, \theta_2 \ \text{ with } \ \theta_2 > 0$$

vs.

$$H_1\colon \ F(x) \neq G((x-\theta_1)/\theta_2) \ \text{ for any } \ \theta_1, \theta_2 \ \text{ with } \ \theta_2 > 0$$

Thus, we are interested in determining whether or not $F(x)$ belongs to a specified scale-location parameter family. The function $G(x)$ is a standardized representative of that family; the nuisance parameters θ_1 and θ_2 are unknown and unspecified.

The test statistic which we will be concerned with has been proposed and analyzed under the null hypothesis by Weiss [41] and is based upon a certain subset of the ordered sample values $Y_1(n) \leq Y_2(n) \leq \ldots \ Y_n(n)$. In this chapter, we determine the asymptotic distribution of Weiss' test under sequences of alternatives which approach the null hypothesis as n tends to infinity.

In section 2.1, we define certain parameters and state certain assumptions which those parameters must satisfy in order to guarantee desirable asymptotic proterties of the test statistic. We also discuss

the concept of asymptotic indistinguishability of sequences of distribution functions and its applications to our testing problem.

In section 2.2, we describe the Weiss test of fit.

In section 2.3, we introduce a class of alternative distributions which approach the null hypothesis as $n \to \infty$. We state assumptions on the manner and rate of approach to the null hypothesis which the sequence of alternatives must satisfy in order to yield a non-trivial power in the limit, and we derive several useful results based on those assumptions.

In section 2.4, we derive the asymptotic power of the test.

In section 2.5, we give examples to illustrate the results of the preceding sections.

2.1 Preliminaries

2.1.1 Definitions and Assumptions

For each n, let p_n, $0 < p_n < 1/2$, be such that np_n is an integer. Further, let K_n and L_n be integers such that

$$K_n L_n = n - 2np_n. \tag{2.1}$$

Define $B_1 = \text{glb}\{x: \ G(x) > 0\}$ and $B_2 = \text{lub}\{x: \ G(x) < 1\}$. ($B_1$ may be $-\infty$; B_2 may be $+\infty$.) Define τ_n as

$$\inf \{g(x): \ G^{-1}(p_n) \leqq x \leqq G^{-1}(1-p_n)\}$$

where $g(x)$ denotes the density corresponding to $G(x)$. We assume the following:

$$\lim_{n\to\infty} p_n = 0 \qquad \lim_{n\to\infty} np_n = \infty \qquad \lim_{n\to\infty} K_n = \infty \tag{2.2}$$

$$\lim_{n\to\infty} \frac{1}{p_n K_n^{1/2}} = 0 \tag{2.3}$$

$$\lim_{n\to\infty} \frac{L_n}{n\tau_n^2} = 0 \tag{2.4}$$

$$\lim_{n\to\infty} \frac{K_n}{n^{1/2}\tau_n^3} = 0 \tag{2.5}$$

$$\lim_{n\to\infty} \frac{n}{L_n^{3/2}\tau_n^2} = 0 \tag{2.6}$$

$g(x) < D < \infty$ for all x in (B_1,B_2) (2.7)

$g'(x)$ exists and $|g'(x)| < D_1 < \infty$ for all x in (B_1,B_2) (2.8)

$|x|g(x) < D_2 < \infty$ for all x in (B_1,B_2) (2.9)

There exist values ζ_1,ζ_2 with $B_1 \leqq \zeta_1 < \zeta_2 \leqq B_2$

$0 < \overline{\zeta} < g(x)$ for all x in (ζ_1,ζ_2) (2.10)

$g(x)$ is nondecreasing in (B_1,ζ_1)

$g(x)$ is nonincreasing in (ζ_2, B_2)

$$\int_{B_1}^{B_2} \left[y\frac{g'(y)}{g(y)}\right]^2 g(y)dy < \infty \tag{2.11}$$

2.1.2 Some Relations Based on the Assumptions

The assumptions (2.1) - (2.11) imply other relations which will also be required. These relations together with the assumptions implying them are summarized in Table 2.1. Verifying these relations is straightforward and will not be done here.

2.1.3 Notation

Throughout the chapter we will use bars over the names of random variables which are used to illustrate various properties or represent certain distributions in order to distinguish them from those random variables directly connected with the problem p_n. Primes will often be used to indicate differentiation of functions; they will also be used to indicate transposes of vectors. Since no vectors will be differentiated, no confusion should arise. In order to eliminate many unnecessary variable names, $o(.)$, $o_p(.)$, and $O(.)$ notation will be used throughout. We write "$x(n)$ is $o(y(n))$" if $x(n)/y(n)$ converges to zero as $n \to \infty$. Similarly, "$x(n)$ is $o_p(y(n))$" if $x(n)/y(n)$ converges to zero in probability as $n \to \infty$; "$x(n)$ is $O(y(n))$" if $x(n)/y(n)$ is finite as $n \to \infty$. As is traditional, we use $\Phi(x)$ and $\phi(x)$ to represent the c.d.f. and p.d.f., respectively, of the standard normal (or $N(0,1)$) distribution. Finally, in order to simplify notation, the dependence on n of $X_1(n),\ldots,X_n(n)$ and $Y_1(n),\ldots,Y_n(n)$ will be suppressed, and we will write $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ instead.

2.1.4 The Joint Asymptotic Normality of an Increasing Number of Order Statistics

The joint asymptotic normality of a fixed set of suitably standardized sample quantiles from a given continous distribution is well known (see e.g. David [8,ch. 9]). The result below is a generalization to the case where: (1) the number of sample quantiles is no longer fixed

TABLE 2.1

IMPLICATIONS OF THE ASSUMPTIONS

Relation	No.	Implied By
$\lim_{n\to\infty} \frac{L_n}{np_n} = 0$	(2.12)	(2.1)&(2.3)
$\lim_{n\to\infty} \frac{n}{L_n^{3/2}} = 0$	(2.13)	(2.6)&(2.7)
$\lim_{n\to\infty} \frac{K_n}{L_n^{1/2}\tau_n} = 0$	(2.14)	(2.6)
$\lim_{n\to\infty} \frac{p_n}{n\tau_n^4} = 0$	(2.15)	(2.4)
$\lim_{n\to\infty} \frac{n}{L_n^2\tau_n^4} = 0$	(2.16)	(2.4)&(2.6)
$\lim_{n\to\infty} \frac{n}{L_n^3\tau_n^6} = 0$	(2.17)	(2.4)&(2.6)
$\lim_{n\to\infty} \frac{n}{L_n^{5/2}\tau_n^4} = 0$	(2.18)	(2.4)&(2.6)
$\int_{B_1}^{B_2} [\frac{g'(x)}{g(x)}]^2 g(x)dx < \infty$	(2.19)	(2.11)
$\int_{B_1}^{B_2} x[\frac{g'(x)}{g(x)}]^2 g(x)dx < \infty$	(2.20)	(2.11)

but is instead allowed to depend on n and slowly increase as n does; and (2) the underlying distribution is replaced by a sequence of underlying distributions, one for each value of n. We require the following definition:

Definition 2.1. Suppose for each n, $F_n(x_1,\ldots,x_{k(n)})$ and $G_n(x_1,\ldots,x_{k(n)})$ are joint cumulative distribution functions for $k(n)$-dimensional random variables. The sequences $\{F_n\}$ and $\{G_n\}$ are Asymptotically Indistinguishable (A.I.) if

$$\lim_{n\to\infty} \left|P_{F_n}\{(X_1,\ldots,X_{k(n)})\ \varepsilon R(n)\} - P_{G_n}\{(X_1,\ldots,X_{k(n)})\ \varepsilon R(n)\}\right| = 0$$

for any sequence of measurable regions $\{R(n)\}$ of $k(n)$-dimensional Euclidean space.

Suppose $f_n(x_1,\ldots,x_{k(n)})$ and $g_n(x_1,\ldots,x_{k(n)})$ are probability density functions corresponding to F_n and G_n respectively. The following condition (Weiss [38]) is necessary and sufficient for Asymptotic Indistinguishability:

$$\frac{f_n(X_1,\ldots,X_{k(n)})}{g_n(X_1,\ldots,X_{k(n)})} \text{ converges stochastically to 1 as } n \to \infty$$

assuming $(X_1,\ldots,X_{k(n)})$ has distribution F_n for each n.

(By symmetry, we can replace F_n by G_n and/or invert the above ratio without affecting the result.) Definition 2.1 defines an equivalence relation, and we will say that the random vectors corresponding to $\{F_n\}$ and $\{G(n)\}$ are asymptotically equivalent when $\{F_n\}$ and $\{G_n\}$ are A.I.

Now, for each n, let $\overline{X}_1,\ldots,\overline{X}_n$ be i.i.d. random variables with absolutely continuous distribution $F_n(x)$. Denote by $\overline{Y}_1,\ldots,\overline{Y}_n$ the corresponding order statistics. Let p_n, K_n, L_n satisfy (2.1) through (2.6) with τ_n replaced by

$$\overline{\tau}_n \equiv \inf\{f_n(x)\colon F_n^{-1}(p_n) \leqq x \leqq F_n^{-1}(1-p_n)\},$$

and assume (2.7) and (2.8) are satisfied for all n with $g(x)$ replaced by $f_n(x)$ and B_1 and B_2 replaced by $\overline{B}_1 \equiv \mathrm{glb}\{x\colon F_n(x) > 0\}$ and $\overline{B}_2 \equiv \mathrm{lub}\{x\colon F_n(x) < 1\}$ respectively. Note that (2.12) - (2.18) also hold. Additionally, assume the following:

$$\lim_{n\to\infty} (n/p_n)^{1/2} f_n(F_n^{-1}(p_n))[\overline{B}_1 - F_n^{-1}(1-p_n)] = -\infty \tag{2.21}$$

and

$$\lim_{n\to\infty} (n/p_n)^{1/2} f_n(F_n^{-1}(1-p_n))[\overline{B}_2 - F_n^{-1}(1-p_n)] = \infty. \tag{2.22}$$

Finally, define $\overline{Z}_j(n)$ as

$$\sqrt{n}\, f_n\left(F_n^{-1}\left(\frac{np_n+(j-1)L_n}{n}\right)\right)\left[\overline{Y}_{np_n+(j-1)L_n} - F_n^{-1}\left(\frac{np_n+(j-1)L_n}{n}\right)\right]$$

for $j = 1,2,\ldots,K_n+1$.

Then Weiss [39] has shown that the distribution of the (K_n+1)-vector $\overline{Z}(n) = (\overline{Z}_1(n),\ldots,\overline{Z}_{K_n+1}(n))$ is asymptotically equivalent to the (K_n+1)-variate normal distribution with mean vector 0 and covariance matrix $\sum(n) = [\sigma_{ij}(n)]$, where

$$\sigma_{ij}(n) = \frac{L_n}{L_n-1}(p_n + (i-1)L_n/n)(1 - p_n - (j-1)L_n/n) \quad (2.23)$$

for $1 \leq i \leq j \leq K_n+1$.

We remark that the result of Weiss [39] is actually proved in more generality than is stated above. In particular, Weiss allows for the possibility that the bounds $\overline{B}_1$ and $\overline{B}_2$ on the support of $F_n(x)$ may depend on n, as might the upper bound D_1 on $f_n(x)$. However, for our purposes these extensions are unnecessary. Also, returning to the assumptions of section 2.1.1, Weiss [41] has shown that assumptions (2.9) and (2.10) together with (2.1) - (2.8) are sufficient to insure that (2.21) and (2.22) hold. Hence, the asymptotic normality result holds with $f_n(x)$ and $F_n(x)$ replaced by $g(x)$ and $G(x)$, respectively, in the definition of $\overline{Z}(n)$.

Reiss [21] has further generalized the above results (in terms of the conditions placed on the "spacing" of the indices of the selected order statistics) and obtained Barry-Esseen type bounds on the rate at which the asymptotic normality takes hold.

Our motivation for considering an increasing subset of order statistics as defined above is the desire to have a manageable distribution theory while, at the same time, not having to sacrifice too much of the information contained in the sample. If we consider a fixed (finite) number of sample quantiles, we have an asymptotic distribution theory which is easy to work with, but, in neglecting the remaining order statistics, we pay a high price in terms of lost information. On the other hand, in [40], Weiss has shown for a simple hypothesis that an increasing subset of order statistics defined as above is asymptotically sufficient for the sample; thus, in the limit, no information is lost. (Weiss proves the result for the "equally spaced" case, i.e. $L_n = np_n$, but his result may

be easily extended to the case $L_n = o(np_n)$ which we consider here.) By asymptotic sufficiency, we mean that using only the statistics $\{Y_{np_n+(i-1)L_n};\ i=1,2,\ldots,K_n+1\}$ together with knowledge of the distribution $G(x)$ governing the population from which the sample was drawn, it is possible to generate a random n-vector with the same asymptotic distribution as the complete vector of order statistics. Furthermore, and of more importance in studying tests of fit, the same results holds if F_n rather than G is the true distribution function for the sample, where $\{F_n\}$ converges to G according to certain regularity conditions.

2.1.5. Some Variances Related to the Multivariate Normal and the Inverse Covariance Matrix $\sum^{-1}(n)$.

Assuming $\overline{Z}(n) = (\overline{Z}_1(n),\ldots,\overline{Z}_{K_n+1}(n))$ is a (K_n+1)-variate normal random variable with mean 0 and covariance matrix $\sum(n)$, (2.23) implies

$$\operatorname{var}(\overline{Z}_1(n)) = \operatorname{var}(\overline{Z}_{K_n+1}(n)) = p_n(1 + o(1)), \tag{2.24}$$

$$\operatorname{var}(\overline{Z}_j(n) - \overline{Z}_{j-1}(n)) = (L_n/n)(1 + o(1)) \tag{2.25}$$

and, for $i < j$,

$$\operatorname{cov}(\overline{Z}_i(n)-\overline{Z}_{i-1}(n),\overline{Z}_j(n)-\overline{Z}_{j-1}(n)) = -\frac{L_n}{L_n-1}(L_n/n)^2. \tag{2.26}$$

Futhermore, for $1 \leq j \leq K_n+1$, $\operatorname{var}(\overline{Z}_j(n))$ is bounded by a finite constant independent of j and n. We will frequently have occasion to refer to the inverse covariance matrix $\sum^{-1}(n)$. It is easily verified that $\sum^{-1}(n)$ is given by $c_n\sum_1^{-1}(n)$, where $c_n = n(L_n-1)L_n^{-2}$ and $\sum_1(n)$ denotes the

$(K_n+1)\times(K_n+1)$ matrix

$$\begin{bmatrix} \frac{L_n}{np_n}+1 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & \frac{L_n}{np_n}+1 \end{bmatrix} . \qquad (2.27)$$

In appendix B, we describe a transformation of the vector $\overline{Z}(n)$ to a vector of i.i.d. standard normal random variables. In particular, we present a nonsingular matrix $B(n)$ such that $B'(n)\Sigma^{-1}(n) = I_{K_n+1}$, the $(K_n+1)\times(K_n+1)$ identity matrix.

2.2 Construction of the Test

Before describing the test statistic for the problem P_n, we first introduce some further simplifying notation. Denote the quantity $G^{-1}(\frac{np_n+(j-1)L_n}{n})$ by $U_j(n)$ and $g(U_j(n))$ by $u_j(n)$ for $j = 1,2,\ldots,K_n+1$. if H_0 is true, then $F(x) = G((x-\theta_1)/\theta_2)$ for all x,

$$F^{-1}(t) = \theta_1 + \theta_2 G^{-1}(t) \qquad (2.28)$$

and

$$f(F^{-1}(t)) = \frac{1}{\theta_2} g(G^{-1}(t)) \qquad (2.29)$$

for all $t \in (0,1)$. Denote $\frac{\sqrt{n}\, u_j(n)}{\theta_2}(Y_{np_n+(j-1)L_n} - \theta_1 - \theta_2 U_j(n))$ by $Z_j(n;\theta_1,\theta_2)$, and let $Z(n;\theta_1,\theta_2),\ldots,Z_{K_n+1}(n;\theta_1,\theta_2))$.

Henceforth assume without loss of generality that $G^{-1}(1/2) = 0$ and $G^{-1}(3/4) - G^{-1}(1/4) = 1$. This can be accomplished by linearly transforming the parameters and has the advantage that the estimators for θ_1 and θ_2 which we will be using will be free of constant terms which depend on $G(x)$.

Finally, let $A_1 = \theta_2 B_1 + \theta_1$ and $A_2 = \theta_2 B_2 + \theta_1$ denote the bounds on the support of $F(x)$ under H_0.

2.2.1 Estimating the Parameters

For each n, we will estimate θ_1 and θ_2 by $\hat{\theta}_1(n)$ and $\hat{\theta}_2(n)$, respectively, where $\hat{\theta}_1(n)$ is the sample median of the sample $(X_1,\ldots,X_n)$, and $\hat{\theta}_2(n)$ is the sample interquartile range (the difference between the third sample quartile and the first sample quartile). Denote $\sqrt{n}(\hat{\theta}_1(n) - \theta_1)$ by $R_1(n)$ and $\sqrt{n}(\hat{\theta}_2) - \theta_2)$ by $R_2(n)$.
Then $R_1(n)$ and $R_2(n)$ are asymptotically jointly normally distributed with zero means and finite variances (see, e.g., David [8,p. 201]). Alternatively, we can write

$$\begin{aligned} \hat{\theta}_1(n) &= \theta_1 + R_1(n)/\sqrt{n} \\ \hat{\theta}_2(n) &= \theta_2 + R_2(n)/\sqrt{n} \quad . \end{aligned} \tag{2.30}$$

We should note that there are more efficient estimators than $\hat{\theta}_1(n)$ and $\hat{\theta}_2(n)$, but the latter have the advantage of simplicity and at the same time are "efficient enough" for asymptotic purposes. That is,

asymptotically we only require that $\hat{\theta}_1(n)$ and $\hat{\theta}_2(n)$ approach θ_1 and θ_2 at a certain rate; in particular, that $\sqrt{n}\,(\hat{\theta}_1(n) - \theta_1)$ and $\sqrt{n}(\hat{\theta}_2(n) - \theta_2)$ are finite with probability one as $n \to \infty$.

2.2.2 The Test Statistic and Critical Region

Under the null hypothesis with θ_1 and θ_2 as the true parameter values, the assumptions (2.1) - (2.10) imply that $Z(n;\theta_1,\theta_2)$ is asymptotically normal with mean 0 and covariance matrix $\sum(n)$. Thus, the quadratic form

$$M(n;\theta_1,\theta_2) \equiv c_n Z'(n;\theta_1,\theta_2){\sum}_1^{-1}(n)Z(n;\theta_1,\theta_2)$$

is asymptotically equivalent to a chi-square random variable with K_n+1 degrees of freedom. Standardizing $M(n;\theta_1,\theta_2)$, we obtain the result that

$$T(n;\theta_1,\theta_2) \equiv \frac{M(n;\theta_1,\theta_2) - (K_n+1)}{\sqrt{2(K_n+1)}}$$

is asymptotically normal with mean 0 and variance 1.

The statistic $T(n;\theta_1,\theta_2)$ is, of course, not suitable for testing the hypothesis of problem P_n since it involves the unknown parameters θ_1 and θ_2. In [41], however, Weiss proves that θ_1 and θ_2 can be replaced by their estimates $\hat{\theta}_1(n)$ and $\hat{\theta}_2(n)$ without changing the asymptotic result. The computable statistic $T(n;\hat{\theta}_1,\hat{\theta}_2)$ is asymptotically $N(0,1)$. (Weiss' proof is done under assumptions (2.4) and (2.7) - (2.11) for specific functions p_n,K_n,L_n, but a review of the proof indicates that (2.1) - (2.3) and (2.5) - (2.6) are sufficient to guarantee the result.) Thus, for fixed $\alpha \in (0,1)$, if we define $c_{\Phi,1-\alpha}$ by

$$\int_{-\infty}^{c_{\Phi,1-\alpha}} \phi(x)dx = 1 - \alpha,$$

a level-α test of H_0 is given by

$$\text{Reject } H_0 \text{ iff } T(n;\hat{\theta}_1,\hat{\theta}_2) > c_{\Phi,1-\alpha}. \tag{2.31}$$

The upper tail of the distribution is chosen as the critical region because $T(n;\hat{\theta}_1,\hat{\theta}_2)$ tends to be larger under alternatives than under H_0. In [41], Weiss also shows that this test is consistent under mild regularity conditions. By this we mean that for any fixed $F(x)$ in the alternative hypothesis, the power of the test (2.31) converges to 1 as n tends to infinity.

2.3 A General Class of Contiguous Alternatives

The remainder of this chapter is concerned with determining the power of the test under sequences of alternatives which approach the null hypothesis as n gets large. Such alternatives are termed "contiguous." By properly specifying the rate of approach, we are able to exhibit a power in the open interval $(\alpha,1)$ and thus provide a basis of comparison between (2.31) and other known tests.

The alternative distributions $H_n(x)$ which we will study have the form

$$H_n(x) = G((x-\theta_1^*)/\theta_2^*) + \varepsilon_n(x) \tag{2.32}$$

where θ_1^* and θ_1^* are given constants, $\theta_2^* > 0$ and where, denoting

$\theta_2^*B_1 + \theta_1^*$ by A_1^* and $\theta_2^*B_2 + \theta_1^*$ by A_2^*, the functions $\varepsilon_n(x)$, which which we will refer to as "disturbance functions," satisfy the following:

$$\varepsilon_n''(x) \text{ exists and is finite for all } x \text{ in } (A_1^*, A_2^*) \tag{2.33}$$

$$-G((x-\theta_1^*)/\theta_2^*) \leqq \varepsilon_n(x) \leqq 1 - G((x-\theta_1^*)/\theta_2^*) \text{ for all } x \text{ in } (A_1^*, A_2^*) \tag{2.34}$$

$$\varepsilon_n(x) = 0 \quad \text{for all } x \notin (A_1^*, A_2^*) \tag{2.35}$$

$$\int_{A_1^*}^{A_2^*} \varepsilon_n'(x)dx = 0 \tag{2.36}$$

$$\varepsilon_n'(x) \geqq -\frac{1}{\theta_2^*} g((x-\theta_1^*)/\theta_2^*) \quad \text{for all } x \text{ in } (A_1^*, A_2^*) \tag{2.37}$$

Condition (2.33) is a regularity condition. Relations (2.34) through (2.37) insure that $H_n(x)$ is in fact a distribution function. In most applications, we will specify $H_n(x)$ rather than $\varepsilon_n(x)$, and hence, (2.34) - (2.37) will automatically be satisfied.

None of the above conditions specifies anything about the rate at which the sequence of alternatives $\{H_n(x)\}$ approaches $G((x-\theta_1^*)/\theta_2^*)$. To insure that the asymptotic power will fall between the level α of the test and 1, we define

$$\begin{aligned} \omega(n) &= \max\{|u_1(n)U_1(n)|, |u_{K_n+1}(n)U_{K_n+1}(n)|\} \\ \nu_1(n) &= \frac{\varepsilon_n(H_n^{-1}(1/2))}{g(0)} \\ \nu_2(n) &= \frac{\varepsilon_n(H_n^{-1}(3/4))}{g(G^{-1}(3/4))} - \frac{\varepsilon_n(H_n^{-1}(1/4))}{g(G^{-1}(1/4))} \end{aligned} \tag{2.38}$$

and add the requirements:

$$\operatorname*{Limsup}_{n\to\infty} \frac{L_n^{1/2}K_n^{1/4}}{p_n^{1/2}}\,\omega(n)|\nu_i(n)| = d_i \qquad i=1,2 \tag{2.39}$$

$$\lim_{n\to\infty} L_n K_n^{1/2} \int_{A_1^*}^{A_2^*} \frac{[\varepsilon_n'(x)]^2}{h_n(x)}\,dx = d_3 \tag{2.40}$$

$$\lim_{n\to\infty} b(n) \sup_{H_n^{-1}(p_n)<x<H_n^{-1}(1-p_n)} \left|\frac{\varepsilon_n''(x)}{h_n(x)}\right| = 0$$

for some increasing sequence $b(n)$ with (2.41)

$$\lim_{n\to\infty} b(n) = \infty, \quad \lim_{n\to\infty} \frac{L_n^{5/4}}{n^{3/4}\tau_n b(n)} = 0$$

where d_1, d_2, d_3 are nonnegative, finite, and not all zero. The motivation for conditions (2.39) and (2.40) will become clear in the derivation of the power of the test. Heuristically, the integral appearing in (2.40) is the "characteristic metric" for the test statistic $T(n;\hat{\theta}_1,\hat{\theta}_2)$. Condition (2.39) requires that the estimators $\hat{\theta}_1(n)$ and $\hat{\theta}_2(n)$ converge sufficiently rapidly in some cases. If $B_1 = -\infty$ and $B_2 = +\infty$, it is easily shown that $p_n^{-1/2}\omega(n)$ converges to zero as $n \to \infty$. In that case, (2.40) implies (2.39) with $d_1 = d_2 = 0$ (see (2.42) below). Condition (2.41) is a regularity condition on the tails of $\varepsilon_n''(x)$ which enables us to approximate complicated sums by much simpler integrals. In most cases, (2.41) will also be implied by (2.40).

<u>Lemma 2.1</u>. <u>If (2.40) holds, then</u>

$$\lim_{n\to\infty} L_n^{1/2} K_n^{1/4} \sup_{A_1^*<x<A_2^*} |\varepsilon_n(x)| \leq d_3^{1/2} \tag{2.42}$$

$$\lim_{n\to\infty} \frac{L_n^{1/2} K_n^{1/4}}{p_n^{1/2}} \varepsilon_n(H_n^{-1}(p_n)) = 0 \tag{2.43}$$

$$\lim_{n\to\infty} \frac{L_n^{1/2} K_n^{1/4}}{p_n^{1/2}} \varepsilon_n(H_n^{-1}(1-p_n)) = 0 \tag{2.44}$$

<u>Proof</u>. For any $x \in (A_1^*, A_2^*)$,

$$L_n^{1/2} K_n^{1/4} |\varepsilon_n(x)| = L_n^{1/2} K_n^{1/4} \left|\int_{A_1^*}^{x} \varepsilon_n'(y)dy\right|$$

$$\leq L_n^{1/2} K_n^{1/4} \int_{A_1^*}^{A_2^*} \frac{|\varepsilon_n'(x)|}{h_n^{1/2}(x)} h_n^{1/2}(x)dx$$

$$\leq \{L_n K_n^{1/2} \int_{A_1^*}^{A_2^*} \frac{[\varepsilon_n'(x)]^2}{h_n(x)} dx\}^{1/2} \{\int_{A_1^*}^{A_2^*} h_n(x)dx\}^{1/2}.$$

By (2.40), the right hand side approaches d_3 as $n \to \infty$, from which (2.42) follows. Using the same method, but over the interval $(A_1^*, H_n^{-1}(p_n))$, we obtain the sharper result (2.43) for the tail behavior of $\varepsilon_n(x)$. We have

$$L_n^{1/2} K_n^{1/4} |\varepsilon_n(H_n^{-1}(p_n))| \leq \{L_n K_n^{1/2} \int_{A_1^*}^{H_n^{-1}(p_n)} \frac{[\varepsilon_n'(x)]^2}{h_n(x)} dx\}^{1/2} \{\int_{A_1^*}^{H_n^{-1}(p_n)} h_n(x)dx\}^{1/2}$$

$$\leq \{L_n K_n^{1/2} \int_{A_1^*}^{H_n^{-1}(p_n)} \frac{[\varepsilon_n'(x)]^2}{h_n(x)} dx\}^{1/2} \; p_n^{1/2}$$

which, after dividing both sides by $p_n^{1/2}$ and taking limits as $n \to \infty$, yields (2.43). Finally, (2.44) follows in exactly the same manner when the interval $(H_n^{-1}(1-p_n), A_2^*)$ is used. This completes the proof of the lemma. QED

Lemma 2.2. If (2.40) holds, then

$$\lim_{n\to\infty} L_n^{1/2} K_n^{1/4} \left| \int_0^1 \frac{\varepsilon_n'(H_n^{-1}(t))}{h_n(H_n^{-1}(t))} \frac{g'(G^{-1}(t))}{g(G^{-1}(t))} G^{-1}(t)dt \right| < \infty.$$

Proof. Using the same method as in the proof of lemma 2.1, we have

$$L_n^{1/2} K_n^{1/4} \int_0^1 \frac{\varepsilon_n'(H_n^{-1}(t))}{h_n(H_n^{-1}(t))} \frac{g'(G^{-1}(t))}{g(G^{-1}(t))} G^{-1}(t)dt \Big| \leq$$

$$\{L_n K_n^{1/2} \int_0^1 [\frac{\varepsilon_n'(H_n^{-1}(t))}{h_n(H_n^{-1}(t))}]^2 dt\}^{1/2} \{\int_0^1 [\frac{g'(G^{-1}(t))}{g(G^{-1}(t))} G^{-1}(t)]^2 dt\}^{1/2}.$$

Making the transformations $x = H_n^{-1}(t)$ in the first integral and $y = G^{-1}(t)$ in the second integral, the right hand side becomes

$$\{L_n K_n^{1/2} \int_{A_1^*}^{A_2^*} \frac{[\varepsilon_n'(x)]^2}{h_n(x)} dx\}^{1/2} \{\int_{A_1^*}^{A_2^*} [y \frac{g'(y)}{g(y)}]^2 dy\}^{1/2}.$$

Taking limits of both sides as $n \to \infty$ and applying (2.11) and (2.40),

the result follows.

2.3.1 An Expansion

In this subsection, we establish the following useful relation:

$$H_n^{-1}(t) = \theta_1^* + \theta_2^* G^{-1}(t) - \theta_2^* \frac{\varepsilon_n(H_n^{-1}(t))}{g(G^{-1}(t))} + \Delta_1(n,t) \qquad (2.45)$$

for $t \in (0,1)$, where $\Delta_1(n,t)/\varepsilon_n(H_n^{-1}(t))$ goes to zero as $n \to \infty$ for all $t \in (0,1)$. Alternatively, we may write (2.45) as

$$\frac{g(G^{-1}(t))}{\theta_2^*}\,(H_n^{-1}(t) - \theta_1^* - \theta_2^* G^{-1}(t)) = -\varepsilon_n(H_n^{-1}(t))\,(1 + o(1)) \qquad (2.45')$$

or, using (2.42),

$$\frac{g(G^{-1}(t))}{\theta_2^*}(H_n^{-1}(t) - \theta_1^* - \theta_2^* G^{-1}(t)) = -\varepsilon_n(H_n^{-1}(t)) + o(L_n^{-1/2}K_n^{-1/4}). \qquad (2.45'')$$

To prove (2.45), let $x = H_n^{-1}(t)$. Then, from the definition of $H_n(x)$,

$$t = H_n(x) = G((x-\theta_1^*)/\theta_2^*) + \varepsilon_n(x) = G((x-\theta_1^*)/\theta_2^*) + \varepsilon_n(H_n^{-1}(t)),$$

so that

$$x = \theta_2^* G^{-1}(t - \varepsilon_n(H_n^{-1}(t))) + \theta_1^*.$$

Expanding $G^{-1}(t - \varepsilon_n(H_n^{-1}(t)))$ in a Taylor series about $G^{-1}(t)$ yields

$$G^{-1}(t - \varepsilon_n(x)) = G^{-1}(t) - \frac{\varepsilon_n(x)}{g(G^{-1}(t))} - \frac{\varepsilon_n^2(x)}{2}\,\frac{g'(G^{-1}(\xi))}{g^3(G^{-1}(\xi))}$$

for some $\xi \in (t, t-\varepsilon_n(H_n^{-1}(t))$. Substituting the above into the expression for x completes the proof.

2.3.2 Asymptotic Normality of Order Statistics and the Asymptotic Distributions of the Estimators under the Alternatives

In order to derive the asymptotic power of the test (2.31) under the sequence of alternatives $\{H_n(x)\}$, we will require the asymptotic normality of the (K_n+1)-vector $Z^*(n) = (Z_1^*(n),\ldots,Z_{K_n+1}^*(n))$, where $Z_j^*(n)$ is given by

$$\sqrt{n}\; h_n\left(H_n^{-1}\left(\frac{np_n+(j-1)L_n}{n}\right)\right)\left[Y_{np_n+(j-1)L_n} - H_n^{-1}\left(\frac{np_n+(j-1)L_n}{n}\right)\right].$$

Let τ_n^* denote $\inf_x h_n(x)$: $H_n^{-1}(p_n) \leq x \leq H_n^{-1}(1-p_n)$. Then the asymptotic normality will follow if (a) relations (2.4) - (2.6) are satisfied with τ_n replaced by τ_n^* for all feasible (p_n,K_n,L_n) combinations, and (b) conditions (2.21) and (2.22) are satisfied with $F_n^{-1}(.)$ and $f_n(.)$ replaced by $H_n^{-1}(.)$ and $h_n(.)$, respectively. A sufficient condition for both (a) and (b) is that τ_n^*/τ_n remains finite as $n \to \infty$. In particular, it is sufficient to show that

$$\lim_{n\to\infty} \theta_2^* \frac{h_n(H_n^{-1}(p_n))}{g(G^{-1}(p_n))} = 1 \qquad (2.46)$$

since the argument for the other tail is the same. Using (2.45'), we have

$$\theta_2^* \frac{h_n(H_n^{-1}(p_n))}{g(G^{-1}(p_n))} = \frac{g(G^{-1}(p_n) - \frac{\varepsilon_n(H_n^{-1}(p_n))}{g(G^{-1}(p_n))}(1+o(1))) + \varepsilon_n'(H_n^{-1}(p_n))}{g(G^{-1}(p_n))}$$

and the right hand side can be written using a Taylor series expansion as

$$1 - \frac{\varepsilon_n(H_n^{-1}(p_n))g'(G^{-1}(p_n))}{[g(G^{-1}(p_n))]^2} + \frac{\varepsilon_n(H_n^{-1}(p_n))}{g(G^{-1}(p_n))} + \text{smaller order terms.}$$

Now, by (2.41),

$$\frac{\varepsilon_n'(H_n^{-1}(p_n))}{g(G^{-1}(p_n))} = \frac{\varepsilon_n'(H_n^{-1}(p_n))}{h_n(H_n^{-1}(p_n))} \frac{h_n(H_n^{-1}(p_n))}{g(G^{-1}(p_n))} = o\left(\frac{h_n(H_n^{-1}(p_n))}{g(G^{-1}(p_n))}\right)$$

as $n \to \infty$. Also, by 2.8 and (2.43),

$$\left|\frac{\varepsilon_n(H_n^{-1}(p_n))g'(G^{-1}(p_n))}{g(G^{-1}(p_n))}\right| \leq \frac{|\varepsilon_n(H_n^{-1}(p_n))| D_1}{\tau_n^2}$$

$$= o\left(\frac{p_n^{1/2}}{L_n^{1/2} K_n^{1/4} \tau_n^2}\right) ,$$

and (2.1) and (2.6) together imply that the latter expression converges to zero as n tends to infinity. Hence, (2.46) is proved. To demonstrate that (2.46) and its counterpart for the other tail imply (2.21) and (2.22), note that (2.21) can be written in this case as

$$\lim_{n\to\infty} \left(\frac{n}{p_n}\right)^{1/2} \frac{h_n(H_n^{-1}(p_n))}{g(G^{-1}(p_n))}[g(G^{-1}(p_n))(B_1-G^{-1}(p_n))+\theta_2^*\varepsilon_n(H_n^{-1}(p_n))(1+o(1))] = -\infty,$$

which is easily seen to be true using (2.43),(2.2), and the fact that (2.21) is satisfied by $G(x)$.

In summary, we have shown that the random vector $Z^*(n)$ is asymptotically equivalent to a (K_n+1)-variate normal random variable with mean vector zero and covariance matrix $\sum(n)$.

A special case of this asymptotic normality yields the fact that $R_1^* = \sqrt{n}(\hat{\theta}_1 - H_n^{-1}(1/2))$ and $R_2^* = \sqrt{n}(\hat{\theta}_2 - H_n^{-1}(3/4) + H_n^{-1}(1/4))$ are asymptotically jointly normal with zero means and finite variances. Using (2.45), recalling that $G^{-1}(1/2) = 0$ and $G^{-1}(3/4) - G^{-1}(1/4) = 1$, and neglecting terms which are asymptotically of smaller order, we can write

$$\begin{aligned} \hat{\theta}_1(n) &= \theta_1^* + R_1^*/\sqrt{n} + \theta_2^*\nu_1(n) \\ \hat{\theta}_2(n) &= \theta_2^* + R_2^*/\sqrt{n} + \theta_2^*\nu_2(n) \end{aligned} \tag{2.47}$$

where $\nu_1(n)$ and $\nu_2(n)$ are defined by (2.38).

2.4 The Asymptotic Power of the Test

In this section, we derive the asymptotic distribution of the test statistic $T(n;\hat{\theta}_1,\hat{\theta}_2)$ under the sequence of alternatives $\{H_n(x)\}$. In particular, we show that $T(n;\hat{\theta}_1,\hat{\theta}_2)$ is asymptotically normally distributed with variance one and finite, nonnegative mean γ_g given explicitly in lemma 2.5. Knowing the value of the mean, the power is then readily computable by referring to tables of the standard normal c.d.f.

Throughout this section we will refer to the random vector $Z^*(n)$ with components $Z_1^*(n),\ldots,Z_{K_n+1}^*(n)$ as defined in subsection 2.3.2. We will denote $H_n^{-1}\left(\frac{np_n+(j-1)L_n}{n}\right)$ by $U_j^*(n)$ and $h_n(U_j^*(n))$ by $u_j^*(n)$. In addition, define

$$\lambda_j(n) = u_j(n)/\theta_2^* u_j^*(n)$$

and let $\Lambda(n)$ denote the $(K_n+1) \times (K_n+1)$ diagonal matrix whose elements λ_{ij} are given by

$$\lambda_{ij}(n) = \begin{cases} \lambda_j(n) & i=j=1,2,\ldots,K_n+1 \\ 0 & i \neq j \end{cases} .$$

Finally, denote the quantity

$$\sqrt{n}\frac{u_j(n)}{\theta_2^*}(U_j^*(n) - \hat{\theta}_1 - \hat{\theta}_2 U_j(n))$$

by $\eta_j(n;\hat{\theta}_1,\hat{\theta}_2)$ and the (K_n+1)-vector $(\eta_1(n;\hat{\theta}_1,\hat{\theta}_2),\ldots,\eta_{K_n+1}(n;\hat{\theta}_1,\hat{\theta}_2))$ by $\eta(n;\hat{\theta}_1,\hat{\theta}_2)$.

Using the above notation and appropriately arranging terms, it is easily verified that the test statistic $T(n;\hat{\theta}_1,\hat{\theta}_2)$, given by

$$[2(K_n+1)]^{-1/2}(c_n Z'(n;\hat{\theta}_1,\hat{\theta}_2)\Sigma_1^{-1}(n)Z(n;\hat{\theta}_1,\hat{\theta}_2) - (K_n+1))$$

can be written as

$$\frac{c_n}{\sqrt{2(K_n+1)}}\frac{\theta_2^{*2}}{\hat{\theta}_2^{\,2}}\left\{\begin{array}{l} Z^{*\prime}(n)\Lambda(n)\Sigma_1^{-1}(n)\Lambda(n)Z^*(n) - c_n^{-1}(K_n+1) \\ + \ 2Z^{*\prime}(n)\Lambda(n)\Sigma_1^{-1}(n)\eta(n;\hat{\theta}_1,\hat{\theta}_2) \\ + \ \eta'(n;\hat{\theta}_1,\hat{\theta}_2)\Sigma_1^{-1}(n)\eta(n;\hat{\theta}_1,\hat{\theta}_2) \end{array}\right\} \tag{2.48}$$

Expanding $\lambda_j(n)$ as in the preceding section, we have

$$\lambda_j(n) = 1 - \frac{\varepsilon_n(U_j^*(n))g'(U_j(n))}{u_j^2(n)} + \delta_1(n)$$

where $\delta_1(n)$ represents smaller order terms. In particular, using the results of the preceding section, $\lambda_j(n) = 1 + O(L_n^{-1/2}K_n^{-1/4}\tau_n^{-2})$ as $n \to \infty$. Hence, $\Lambda(n)$ can be written as $I_{K_n+1} + \Delta_{K_n+1}$, where Δ_{K_n+1} is a diagonal matrix all of whose diagonal entries are $O(L_n^{-1/2}K_n^{-1/4}\tau_n^{-2})$.

Using the above representation of $\Lambda(n)$ together with the definition (2.27) of $\Sigma_1^{-1}(n)$ and the expansion (2.45'), it is straightforward to verify the following result expressing the test statistic in a form whose asymptotic properties can be investigated.

<u>Lemma 2.3.</u> <u>Under the assymptions of section 2.3, the test statistic</u> $T(n;\hat{\theta}_1,\hat{\theta}_2)$ <u>is given by</u>

$$\frac{c_n}{\sqrt{2(K_n+1)}}\frac{\theta_2^{*2}}{\hat{\theta}_2^{\,2}}\{W^*(n)[1+O(L_n^{-1}K_n^{-1/2}\tau_n^{-4})] - c_n^{-1}(K_n+1) -Q(n)[1+O(L_n^{-1/2}K_n^{-1/4}\tau_n^{-2})] + V(n)]\} \tag{2.49}$$

where $W^*(n)$ is given by

$$Z^{*\prime}(n)\Sigma_1^{-1}(n)Z^*(n), \tag{2.50}$$

$Q(n)$ is the sum of the following six expressions:

$$\frac{L_n}{n^{1/2}p_n}\left\{\begin{array}{l}[u_1(n)(\nu_1(n)+\nu_2(n)) + \varepsilon_n(U_1^*(n))(1+o(1))]Z_1^*(n) + \\ \\ [u_{K_n+1}(n)(\nu_1(n)+\nu_2(n)) + \varepsilon_n(U_{K_n+1}^*(n))(1+o(1))]Z_{K_n+1}^*(n)\end{array}\right\} \tag{2.51}$$

$$\frac{L_n}{np_n}\left\{\frac{u_1(n)}{\theta_2^*}(R_1^* + R_2^*U_1(n))Z_1^*(n) + \frac{u_{K_n+1}(n)}{\theta_2^*}(R_1^* + R_2^*U_{K_n+1}(n))Z_{K_n+1}^*(n)\right\} \tag{2.52}$$

$$\sqrt{n}\sum_{j=1}^{K_n}[\varepsilon_n(U_{j+1}^*(n)) - \varepsilon_n(U_j^*(n))](1+o(1))(Z_{j+1}^*(n) - Z_j^*(n)) \tag{2.53}$$

$$\sqrt{n}\nu_1(n)\sum_{j=1}^{K_n}(u_{j+1}(n) - u_j(n))(Z_{j+1}^*(n) - Z_j^*(n)) \tag{2.54}$$

$$n\nu_2(n)\sum_{j=1}^{K_n}(u_{j+1}(n)U_{j+1}(n) - u_j(n)U_j(n))(Z_{j+1}^*(n) - Z_j^*(n)) \tag{2.55}$$

$$\frac{1}{\theta_2^*}\sum_{j=1}^{K_n}[(u_{j+1}(n) - u_j(n))R_1^* + (u_{j+1}(n)U_{j+1}(n) - u_j(n)U_j(n))R_2^*] \times (Z_{j+1}^*(n) - Z_j^*(n)) \tag{2.56}$$

and $V(n)$ is given by the sum of the following four expressions:

$$\frac{L_n}{p_n}\left\{\begin{array}{l}[\varepsilon_n(U_1^*(n))(1+o(1)) + u_1(n)(\nu_1(n) + \nu_2(n)U_1(n))]^2 \quad + \\ [\varepsilon_n(U_{K_n+1}^*(n))(1+o(1)) + u_{K_n+1}(n)(\nu_1(n) + \nu_2(n)U_{K_n+1}(n))]^2\end{array}\right\} \tag{2.57}$$

$$\frac{2L_n}{\theta_2^* \, np_n}\left\{\begin{array}{l}[\varepsilon_n(U_1^*(n))(1+o(1)) + u_1(n)(\nu_1(n) + \nu_2(n)U_1(n))] \times \\ \qquad [u_1(n)(R_1^* + R_2^*U_1(n))] \quad + \\ [\varepsilon_n(U_{K_n+1}^*(n))(1+o(1)) + u_{K_n+1}(n)(\nu_1(n) + \nu_2(n)U_{K_n+1}(n))] \times \\ \qquad [u_{K_n+1}(n)(R_1^* + R_2^*U_{K_n+1}(n))]\end{array}\right\} \tag{2.58}$$

$$\frac{L_n}{\theta_2^{*2} np_n}\{[u_1(n)(R_1^* + R_2^*U_1(n))]^2 + [u_{K_n+1}(n)(R_1^* + R_2^*U_{K_n+1}(n))]^2\} \tag{2.59}$$

$$n\sum_{j=1}^{K_n}\left\{\begin{array}{l}[\varepsilon_n(U_{j+1}^*(n))(1+o(1)) + u_{j+1}(n)(\nu_1(n) + \dfrac{R_1^*}{\theta_2^*\sqrt{n}} + (\nu_2(n) + \dfrac{R_2^*}{\theta_2^*\sqrt{n}})U_{j+1}(n))] \\ - [\varepsilon_n(U_j^*(n))(1+o(1)) + u_j(n)(\nu_1(n) + \dfrac{R_1^*}{\theta_2^*\sqrt{n}} + (\nu_2(n) + \dfrac{R_2^*}{\theta_2^*\sqrt{n}})U_j(n))]\end{array}\right\}^2 . \tag{2.60}$$

From section 2.3.2, we know that $c_n W^*(n)$ is asymptotically equivalent to a chi-square random variable with (K_n+1) degrees of freedom. The next two lemmas establish the limiting forms for $Q(n)$ and $V(n)$. First note, however, that

$$c_n(K_n+1)^{-1/2} = \frac{n(L_n-1)}{L_n^2(K_n+1)^{1/2}}$$

can be replaced by $K_n^{1/2}$ for all asymptotic calculations.

Lemma 2.4. The quantity $K_n^{1/2}Q(n)$ converges stochastically to zero as n tends to infinity.

Proof. We shall consider each of the terms (2.51) through (2.56) separately. Recall that for asymptotic purposes we can consider $Z_1^*(n),\ldots,$ $Z_{K_n+1}^*(n)$ to be jointly normally distributed as described above.

(i) Using (2.7), (2.24) and (2.42) - (2.44), the quantity in expression (2.51) is asymptotically normal with mean zero and variance of order

$$\frac{L_n^2}{np_n^2}\left(\frac{1}{L_n^{1/2}K_n^{1/4}}\right)^2 p_n = \frac{L_n}{np_nK_n^{1/2}} .$$

Hence, $K_n^{1/2}$ (expression (2.51)) is asymptotically normal with mean zero and variance of order

$$\frac{L_nK_n^{1/2}}{np_n},$$

which is $O(K_n^{-1/2}p_n^{-1})$ as n tends to infinity. Applying assymption (2.3), $K_n^{1/2}$ (expression (2.51)) converges stochastically to zero as $n \to \infty$.

(ii) Using assumptions (2.7) and (2.9), the term in braces in expression (2.52) is finite with probability one. Thus, since $K_n^{1/2}(L_n/np_n)$ converges to zero as $n \to \infty$ as noted in (i) above, $K_n^{1/2}$ (expression (2.52)) converges stochastically to zero as n increases.

(iii) Let $S_1(n)$ denote

$$\sum_{j=1}^{K_n} [\varepsilon_n(U_{j+1}^*(n)) - \varepsilon_n(U_j^*(n))]^2$$

and let $S_2(n)$ denote

$$2\sum_{i=1}^{K_n-1}\sum_{j=i+1}^{K_n}[\varepsilon_n(U^*_{j+1}(n)) - \varepsilon_n(U^*_j(n))][\varepsilon_n(U^*_{i+1}(n) - \varepsilon_n(U^*_i(n)).$$

Then, using (2.25) and (2.26), expression (2.53) is asymptotically normal with mean zero and variance given by

$$L_nS_1(n)(1+o(1)) - \frac{L_n^2}{n}S_2(n)(1+o(1)).$$

Note that

$$S_1(n) = (L_n/n)\{\int_{A_1^*}^{A_2^*}\frac{[\varepsilon_n'(x)]^2}{h_n(x)}dx + o(1)\}$$

and

$$S_2(n) = -S_1(n) + o(p_nL_n^{-1}K_n^{-1/2})$$

where the latter expression follows from (2.43), (2.44) and the fact that

$$S_1(n) + S_2(n) = \{\sum_{j=1}^{K_n}[\varepsilon_n(U^*_{j+1}(n)) - \varepsilon_n(U^*_j(n))]\}^2$$

$$= \{\varepsilon_n(U^*_{K_n+1}(n)) - \varepsilon_n(U^*_1(n))\}^2 \quad .$$

(In the remainder of this proof, and in the proof of the following lemma, we will frequently pass from sums to integrals in the above manner, Assumption (2.41) insures that the errors of such approximations are always

of smaller order than the integrals themselves. The formal procedure is carried out in detail in appendix C, and upper bounds on the errors involved are obtained.)

Thus, by (2.40), $L_n S_1(n) = O(L_n^{3/2} n^{-3/2})$ and $(L_n^2/n)S_2(n) = O(L_n^{5/2} n^{-5/2})$ as $n \to \infty$. In particular, $(L_n/n)^{-3/2}|L_n S_1(n)|$ and $(L_n/n)^{-3/2}|(L_n^2/n)S_2(n)|$ both remain bounded as n increases. Thus, since (L_n/n) converges to zero, $K_n^{1/2}$ (expression (2.53)) converges stochastically to zero as $n \to \infty$.

(iv) Let $S_3(n)$ denote

$$\sum_{j=1}^{K_n} [u_{j+1}(n) - u_j(n)]^2$$

and let $S_4(n)$ denote

$$2 \sum_{i=1}^{K_n - 1} \sum_{j=i+1}^{K_n} [u_{i+1}(n) - u_i(n)][u_{j+1}(n) - u_j(n)].$$

Then, in asymptotic probabity calculations we can assume that expression (2.54) is normally distributed with mean zero and variance given by

$$n v_1^2(n)\{\frac{L_n}{n} S_3(n)(1+o(1)) + \frac{L_n^2}{n^2} S_4(n)(1+o(1))\}.$$

Note that

$$S_3(n) = (L_n/n)\{\int_{B_1}^{B_2} [\frac{g'(x)}{g(x)}]^2 g(x)dx + o(1)\}$$

and

$$S_4(n) = -S_3(n) + [u_{K_n+1}(n) - u_1(n)]^2.$$

Thus, (2.19) implies that $S_3(n) = O(L_n/n)$, $S_4(n)$ is $o(1)$, and, using (2.42) of lemma 2.1, both $L_n\nu_1^2(n)S_3(n)$ and $(L_n^2/n)\nu_1^2(n)S_4(n)$ are $O(L_n^{3/2}n^{-3/2})$. Thus, arguing as in (ii) above, $K_n^{1/2}$ (expression (2.54)) converges stochastically to zero as $n \to \infty$.

(v) Let $S_5(n)$ denote

$$\sum_{j=1}^{K_n} [u_{j+1}(n)U_{j+1}(n) - u_j(n)U_j(n)]^2$$

and let $S_6(n)$ denote

$$2\sum_{i=1}^{K_n-1}\sum_{j=i+1}^{K_n} [u_{i+1}(n)U_{i+1}(n) - u_i(n)U_i(n)][u_{j+1}(n)U_{j+1}(n) - u_j(n)U_j(n)].$$

Note that

$$S_5(n) = (L_n/n)\{\int_{B_1}^{B_2} [1 + x\frac{g'(x)}{g(x)}]^2 g(x)dx + o(1)\}$$

and

$$S_6(n) = -S_5(n) + [u_{K_n+1}(n)U_{K_n+1}(n) - u_1(n)U_1(n)]^2.$$

From (2.9) and (2.11), $S_5(n)$ is $O(L_n/n)$ and $S_6(n)$ is $O(1)$ as $n \to \infty$. Hence, arguing as in (iv) above, we have the result that $K_n^{1/2}$ (expression (2.55)) converges stochastically to zero as n increases.

(vi) Let $S_7(n)$ denote

$$\sum_{j=1}^{K_n} [u_{j+1}(n) - u_j(n)][Z^*_{j+1}(n) - Z^*_j(n)]$$

and let $S_8(n)$ denote

$$\sum_{j=1}^{K_n} [u_{j+1}(n)U_{j+1}(n) - u_j(n)][Z^*_{j+1}(n) - Z^*_j(n)].$$

Then, expression (2.56) can be written as

$$\frac{1}{\theta^*_2}[R^*_1 S_7(n) + R^*_2 S_8(n)].$$

Imbedded in part (iv) above is the fact that $K_n^{1/2} S_7(n)$ converges stochastically to zero as $n \to \infty$. Similarly, in part (v) it is shown that $K_n^{1/2} S_8(n)$ converges stochastically to zero as n increases. Hence, $K_n^{1/2}$ (expression (2.56)) converges stochastically to zero as $n \to \infty$.

Combining (i) through (vi) above, we have shown that $K_n^{1/2} Q(n)$ converges stochastically to zero as n tends to infinity. This completes the proof of lemma 2.4. QED

Define ν^*_1 and ν^*_2 by

$$\nu^*_1 = \lim_{n\to\infty} L_n^{1/2} K_n^{1/4} \nu_1(n)$$

$$\nu^*_2 = \lim_{n\to\infty} L_n^{1/2} K_n^{1/4} \nu_2(n) \quad .$$

(We assume here that the above limits exist, as they will in most reasonable cases. If they do not, due to an oscillating disturbance

function, we may replace ν_1^* and ν_2^* by the appropriate Limsup or Liminf in lemma 2.5 below to obtain upper and lower non-trivial bounds on the asymptotic power of the test. Relation (2.42) of lemma 2.1 insures that the above quantities remain finite as $n \to \infty$.)

Lemma 2.5. *As $n \to \infty$, the quantity $K_n^{1/2}V(n)$ converges stochastically to the constant γ_g given by the sum of the following five expressions:*

$$\lim_{n\to\infty} \frac{L_n K_n^{1/2}}{p_n}\{[u_1(n)(\nu_1(n) + \nu_2(n)U_1(n))]^2 + [u_{K_n+1}(n)(\nu_1(n) + \nu_2(n)U_{K_n+1}(n))]^2\} \tag{2.61}$$

$$d_3 + \nu_1^{*2}\int_{B_1}^{B_2}[\frac{g'(y)}{g(y)}]^2 g(y)dy + 2\nu_1^*\nu_2^*\{\int_{B_1}^{B_2}[\frac{g'(y)}{g(y)}]^2 yg(y)dy + \int_{B_1}^{B_2} g'(y)dy\} \tag{2.62}$$

$$2\nu_1^* \lim_{n\to\infty} L_n^{1/2}K_n^{1/4}\int_{p_n}^{1-p_n} \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}\,\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}\,dx \tag{2.63}$$

$$2\nu_2^* \lim_{n\to\infty} L_n^{1/2}K_n^{1/4}\int_{p_n}^{1-p_n} \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}\,\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}\,G^{-1}(x)dx \tag{2.64}$$

$$\nu_2^{*2}[1 + 2\int_{B_1}^{B_2} yg'(y)dy + \int_{B_1}^{B_2}[y\frac{g'(y)}{g(y)}]^2 g(y)dy]. \tag{2.65}$$

Proof. We consider each of the expressions (2.57) through (2.60) separately.

(i) From (2.43) of lemma 2.1, the quantity

$$\frac{L_n K_n^{1/2}}{p_n}\,\varepsilon_n^2(U_1^*(n))$$

converges to zero as n tends to infinity. Further, we can write

$$\frac{L_n K_n^{1/2}}{p_n} \varepsilon_n^2(U_1^*(n))u_1(n)\nu_1(n) = \frac{L_n^{1/2} K_n^{1/4}}{p_n^{1/2}} \varepsilon_n(U_1^*(n)) \frac{L_n^{1/2} K_n^{1/4}}{p_n^{1/2}} u_1(n)\nu_1(n)$$

and the right hand expression converges to zero by (2.43) and (2.39). In a similar fashion, we can show that the contributions to γ_g of all terms in expression (2.57) which involve $\varepsilon_n(U_1^*(n))$ are asymptotically negligible. Using (2.44) of lemma 2.1, the same conclusion follows for all terms in (2.57) which involve $\varepsilon_n(U_{K_n+1}^*(n))$. Eliminating those terms, we see that $K_n^{1/2}$ (expression (2.57)) is equivalent to (2.61).

(ii) By assumption (2.9), the quantities $u_1(n)[R_1^* + R_2^* U_1(n)]$ and $u_{K_n+1}(n)[R_1^* + R_2^* U_{K_n+1}(n)]$ are asymptotically normally distributed with mean 0 and finite variances. It follows from this fact, the fact implied by (2.3) that

$$\frac{L_n K_n^{1/2}}{n^{1/2} p_n} = o\left(\frac{L_n^{1/2} K_n^{1/4}}{p_n^{1/2}}\right)$$

and the relations (2.43) and (2.44) that the terms involving $\varepsilon_n(U_1^*(n))$ and $\varepsilon_n(U_{K_n+1}^*(n))$ in expression (2.58) approach zero stochastically as n increases. Neglecting those terms, we can write $K_n^{1/2}$ (expression (2.58)) as

$$\frac{L_n^{1/2} K_n^{1/4}}{p_n^{1/2}}[u_1(n)(\nu_1(n) + \nu_2(n)U_1(n))] \frac{L_n^{1/2} K_n^{1/4}}{p_n^{1/2}}[u_1(n)(R_1^* + R_2^* U_1(n))] +$$

$$\frac{L_n^{1/2}K_n^{1/4}}{p_n^{1/2}}[u_{K_n+1}(n)(\nu_1(n) + \nu_2(n)U_{K_n+1}(n))] - \frac{L_n^{1/2}K_n^{1/4}}{p_n^{1/2}}[u_{K_n+1}(n)(R_1^* + R_2^* U_{K_n+1}(n))].$$

Applying (2.39), we see that the above expression converges stochastically to zero as $n \to \infty$.

(iii) Using (2.3), (2.7), and (2.9), it is immediate that $K_n^{1/2}$ (expression (2.59)) converges stochastically to zero as $n \to \infty$.

(iv) Let $S_9(n)$ denote the sum in expression (2.60). Then $S_9(n)$ can be written as

$$(L_n/n)\{ \int_{p_n}^{1-p_n} [\frac{d}{dx} J_n(x)]^2 \, dx + o(L_n^{-1/2}n^{-1/2})\}$$

where $J_n(x)$ is defined as

$$\varepsilon_n(H_n^{-1}(x)) + g(G^{-1}(x))[\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}} + (\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})G^{-1}(x)]$$

and the $o(.)$ term follows from (2.41) as in Appendix C. Expanding $\frac{d}{dx} J_n(x)$, we obtain

$$\frac{d}{dx} J_n(x) = \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))} + \frac{g'(G^{-1}(x))}{g(G^{-1}(x))}[\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}} + (\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})G^{-1}(x)]$$

$$+ \nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}} .$$

Hence, $S_9(n)$ becomes

$$(L_n/n)\{\int_{p_n}^{1-p_n}[\frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}]^2dx + (\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}})\int_{p_n}^{1-p_n}[\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}]^2dx$$

$$+ 2(\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}})(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})\int_{p_n}^{1-p_n}[\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}]^2 G^{-1}(x)dx$$

$$+ (\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})^2\int_{p_n}^{1-p_n}[\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}G^{-1}(x)]^2dx + (\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})^2$$

$$+ 2(\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}})\int_{p_n}^{1-p_n}\frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}dx$$

$$+ 2(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})\int_{p_n}^{1-p_n}\frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}G^{-1}(x)\,dx$$

$$+ 2(\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}})(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})\int_{p_n}^{1-p_n}\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}dx$$

$$+ 2(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})^2\int_{p_n}^{1-p_n}\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}dx$$

$$+ 2(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}})\int_{p_n}^{1-p_n}\frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}dx + o(L_n^{-1/2}n^{-1/2})\}.$$

Combining terms, changing variables of integration, and noting that by (2.43) and (2.44),

$$\int_{p_n}^{1-p_n} \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}\,dx = o(L_n^{-1/2}K_n^{-1/4})$$

we obtain

$$S_9(n) = \frac{L_n}{n}\Big\{\int_{A_1^*}^{A_2^*} \frac{[\varepsilon_n'(y)]^2}{h_n(y)}dy + \Big(\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}}\Big)\int_{B_1}^{B_2} \Big[\frac{g'(y)}{g(y)}\Big]^2 g(y)dy$$

$$+ 2\Big(\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}}\Big)\Big(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}}\Big)\Big(\int_{B_1}^{B_2} \Big[\frac{g'(y)}{g(y)}\Big]^2 yg(y)dy + \int_{B_1}^{B_2} g'(y)dy\Big)$$

$$+ 2\Big(\nu_1(n) + \frac{R_1^*}{\theta_2^* n^{1/2}}\Big)\int_{p_n}^{1-p_n} \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}\,\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}\,dx$$

$$+ 2\Big(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}}\Big)\int_{p_n}^{1-p_n} \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))}\,\frac{g'(G^{-1}(x))}{g(G^{-1}(x))}\,G^{-1}(x)dx$$

$$+ \Big(\nu_2(n) + \frac{R_2^*}{\theta_2^* n^{1/2}}\Big)^2\Big[1 + 2\int_{B_1}^{B_2} yg'(y)dy + \int_{B_1}^{B_2}\Big[y\frac{g'(y)}{g(y)}\Big]^2 g(y)dy\Big]$$

$$+ o(L_n^{-1/2}n^{-1/2})\Big\}.$$

Now let $\tilde{\gamma}$ denote the quantity

$$d_3 + \nu_1^{*2}\int_{B_1}^{B_2}\Big[\frac{g'(y)}{g(y)}\Big]^2 g(y)dy + 2\nu_1^*\nu_2^*\Big(\int_{B_1}^{B_2}\Big[\frac{g'(y)}{g(y)}\Big]^2 yg(y)dy + \int_{B_1}^{B_2} g'(y)dy\Big)$$

$$+ 2\nu_1^* \lim_{n\to\infty} L_n^{1/2} K_n^{1/4} \int_{p_n}^{1-p_n} \frac{\varepsilon_n(H_n^{-1}(x))}{h_n(H_n^{-1}(x))} \frac{g'(G^{-1}(x))}{g(G^{-1}(x))} dx$$

$$+ 2\nu_2^* \lim_{n\to\infty} L_n^{1/2} K_n^{1/4} \int_{p_n}^{1-p_n} \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))} \frac{g'(G^{-1}(x))}{g(G^{-1}(x))} G^{-1}(x)dx$$

$$+ \nu_2^{*2}[1 + 2\int_{B_1}^{B_2} yg'(y) + \int_{B_1}^{B_2} [y\frac{g'(y)}{g(y)}]^2 g(y)dy].$$

Then, since (expression (2.60) = $nS_9(n)$, we can use (2.1) and (2.4) to argue that $K_n^{1/2} n\{$products involving R_1^* and R_2^* in $S_9(n)\}$ converges stochastically to zero, and hence, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{|K_n^{1/2}(\text{expression (2.60)}) - \tilde{\gamma}| > \varepsilon\} = 0$$

Combining (i) through (iv) above completes the proof of lemma 2.5. QED

The asymptotic power of the test (2.31) now follows from lemmas 2.3, 2.4, and 2.5, and is given in the following theorem:

<u>Theorem 2.1</u>. <u>The asymptotic power of the test (2.31) is given by</u>

$$\Phi(\frac{1}{\sqrt{2}} \gamma_g + c_{\Phi,\alpha}) \tag{2.66}$$

<u>as $n \to \infty$</u>.

<u>Proof</u>. First note that

$$\frac{\theta_2^{*2}}{\hat{\theta}_2^{\,2}} = \theta_2^{*2}[\theta_2^{*2} + R_2^*/\sqrt{n} + \nu_2(n)]^{-2}$$

$$= [1 + \frac{R_2^*/\sqrt{n} + \nu_2(n)}{\theta_2^*}]^{-2}$$

and the latter expression converges stochastically to one as $n \to \infty$. From lemma 2.3, we can therefore write the test statistic $T(n;\hat{\theta}_1,\hat{\theta}_2)$ as

$$[1 + \frac{R_2^*/\sqrt{n} + \nu_2(n)}{\theta_2^*}]^{-2}[\frac{c_n W^*(n) - (K_n+1)}{\sqrt{2(K_n+1)}}]\ (1 + o(1)) \ +$$

$$[1 + \frac{R_2^*/\sqrt{n} + \nu_2(n)}{\theta_2^*}]^{-2}\ \frac{Q(n)}{\sqrt{2}}\ (1+o(1)) + [1 + \frac{R_2^*/\sqrt{n} + \nu_2(n)}{\theta_2^*}]^{-2}\ \frac{V(n)}{\sqrt{2}} \ .$$

Using lemmas 2.4 and 2.5, we have

$$T(n;\hat{\theta}_1,\hat{\theta}_2) = \frac{c_n W^*(n) - (K_n+1)}{\sqrt{2(K_n+1)}} + \gamma_g/\sqrt{2} + o_p(1)$$

as $n \to \infty$. Recalling that $c_n W^*(n)$ is asymptotically equivalent to a chi-square random variable with (K_n+1) degrees of freedom completes the proof of theorem 2.1. QED

2.4.1 Remarks

(1) In theory, the asymptotic power of the test immediately obtainable from tables of $\Phi(x)$. The quantity γ_g, however, appears quite formidable to compute, particularly because of the terms (2.63) and (2.64). Since $G(x)$ is usually easier to work with than $H_n(x)$, the following relation

often helps simplify the computation:

$$\frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))} = \frac{\theta_2^*\varepsilon_n'(\theta_1^* + \theta_2^* G^{-1}(x))}{g(G^{-1}(x))}(1+o(1)) \quad \text{for } p_n < x < 1-p_n. \qquad (2.67)$$

(2) None of the conditions imposed on the distrubance function $\varepsilon_n(x)$ specifically required that $H_n(x)$ lie outside the region encompassed by the null hypothesis. In particular, suppose we take

$$H_n(x) = G\left(\frac{x - \theta_1^{(n)}}{\theta_2^{(n)}}\right)$$

where $\theta_1^{(n)} \to \theta_1^*$ and $\theta_2^{(n)} \to \theta_2^*$ as $n \to \infty$. For our definition of alternatives to be consistent, we should obtain a power equal to α for the α-level test (2.31) when $\{H_n(x)\}$ is the underlying sequence of distributions.

First suppose that the assumptions of section 2.3 are satisfied, i.e.

$$\{\varepsilon_n(x) = G\left(\frac{x-\theta_1^{(n)}}{\theta_2^{(n)}}\right) - G\left(\frac{x-\theta_1^*}{\theta_2^*}\right)\}$$

satisfies (2.38) through (2.41). Then it follows that the constant γ_g equals zero and the power is given by $\Phi(c_{\Phi,\alpha}) = \alpha$ as desired. To show this, we note that for this case we have $H_n^{-1}(t) = \theta_1^{(n)} + \theta_2^{(n)}G^{-1}(t)$, and form (2.45''),

$$\varepsilon_n(H_n^{-1}(t)) = -\frac{g(G^{-1}(t))}{\theta_2^*}[(\theta_1^{(n)}-\theta_1^*) + (\theta_2^{(n)}-\theta_2^*)G^{-1}(t)] + o(L_n^{-1/2}K_n^{-1/4}).$$

Differentiating the above expression with respect to t yields

$$\frac{\varepsilon_n'(H_n^{-1}(t))}{h_n(H_n^{-1}(t))} = -\left[\frac{\theta_1^{(n)}-\theta_2^*}{\theta_2^*}\right]\frac{g'(G^{-1}(t))}{g(G^{-1}(t))} - \left[\frac{\theta_2^{(n)}-\theta_2^*}{\theta^*_2}\right]\left[1 + \frac{g'(G^{-1}(t))}{g(G^{-1}(t))}G^{-1}(t)\right].$$

Also,

$$\nu_1(n) = \frac{\varepsilon_n(H_n^{-1}(1/2))}{g(0)} = \frac{\theta_1^{(n)}-\theta_1^*}{\theta_2^*} + o(L_n^{-1/2}K_n^{-1/4})$$

and

$$\nu_2(n) = \frac{\theta_2^{(n)}-\theta_2^*}{\theta_2^*} + o(L_n^{-1/2}K_n^{-1/4})$$

so that

$$\nu_1^* = \lim_{n\to\infty} L_n^{1/2}K_n^{1/4}\left(\frac{\theta_1^{(n)}-\theta_1^*}{\theta_2^*}\right) \quad \text{and} \quad \nu_2^* = \lim_{n\to\infty} L_n^{1/2}K_n^{1/4}\left(\frac{\theta_2^{(n)}-\theta_2^*}{\theta_2^*}\right) .$$

In addition,

$$\varepsilon_n(U_1^*(n)) = \varepsilon_n(H_n^{-1}(p_n)) = -u_1(n)[\nu_1(n) + \nu_2(n)U_1(n)] + o(L_n^{-1/2}K_n^{-1/4}p_n),$$

and similarly,

$$\varepsilon_n(H_n^{-1}(1-p_n)) = -u_{K_n+1}(n)[\nu_1(n) + \nu_2(n)U_{K_n+1}(n)] + o(L_n^{-1/2}K_n^{-1/4}p_n).$$

Finally, note that in this case the assumption (2.36) becomes

$$0 = \frac{\theta_1^{(n)}-\theta_1^*}{\theta_2^*}\int_{B_1}^{B_2} g'(y)dy + \frac{\theta_2^{(n)}-\theta_2^*}{\theta_2^*}\left[\int_{B_1}^{B_2} yg'(y)dy + 1\right],$$

so that

$$\nu_1^* \int_{B_1}^{B_2} g'(y)dy + \nu_2^*\left[\int_{B_1}^{B_2} gy'(y)dy + 1\right] = 0.$$

Using the above relations in the expressions (2.61) - (2.65) yields immediately the desired result that $\gamma_g = 0$.

For the sequence $\{H_n(x)\}$ as defined here, it is of course not necessary that the assumptions of section 2.3 be satisfied in order that the test (2.31) have power equal to α. In fact, we have

$$\eta_j(n;\hat{\theta}_1,\hat{\theta}_2) = \sqrt{n}\,\frac{u_j(n)}{\theta_2^*}[U_j^*(n) - \hat{\theta}_1 - \hat{\theta}_2 U_j(n)]$$

$$= \sqrt{n}\,\frac{u_j(n)}{\theta_2^*}\{[\theta_1^{(n)}+\theta_2^{(n)}U_j(n)]-[\theta_1^{(n)} + \frac{R_1^*}{\sqrt{n}}]-[\theta_2^{(n)} + \frac{R_2^*}{\sqrt{n}}]U_j(n)$$

$$= \frac{u_j(n)}{\theta_2^*}(R_1^* + R_2^*),$$

and it is easily shown that $K_n^{1/2}\eta'(n;\hat{\theta}_1,\hat{\theta}_2)\sum_n^{-1}\eta(n;\hat{\theta}_1,\hat{\theta}_2)$ converges to zero in probability as $n \to \infty$ regardless of the rate at which $\theta_1^{(n)}$ and $\theta_2^{(n)}$ converge to θ_1^* and θ_2^*.

(3) In the preceding sections we defined alternatives relative to a predetermined procedure in order to achieve a power level between α and 1. In particular, the functions L_n and K_n were first specified and then the conditions (2.38) - (2.41) on the alternatives were imposed. In practice, however, the true underlying distribution is unknown but fixed, and it is L_n and K_n which must be chosen. To this end, the form of the power function indicates that choosing L_n as large as possible would

lead to the greatest power for fixed (large) n when the underlying distribution belongs to the alternative hypothesis. The quantity γ_g can be written as $\lim_{n\to\infty} L_n K_n^{1/2} \bar{\gamma}_g(n)$, where $\bar{\gamma}_g(n)$ is some nonnegative function of the true distribution as well as the null hypothesis. Suppose we take $\bar{L}_n = ML_n, \bar{K}_n = (1/M)K_n$ for any positive constant M greater than one. Then it is easy to see that assumptions (2.1) through (2.6) still hold (i.e. $\bar{L}_n$ and $\bar{K}_n$ are feasible), but now γ_g has been increased to $\gamma_g^* = M\gamma_g$, with a corresponding increase in the asymptotic power of the test.

There are, however, additional considerations in choosing L_n. The interpretation of L_n is that of the "spacing" between the successive sample quantiles which are used in computing the test statistic $T(n;\hat{\theta}_1,\hat{\theta}_2)$ (here we mean "spacing" in terms of the indices of the order statistics, not the values which they assume). The larger L_n is made (for fixed n), the fewer the number (K_n+1) of sample quantiles which are used. For "moderate" n, therefore, taking L_n large (say n^δ, where δ is close to 1) may result in an inferior test since very few quantiles will actually contribute to the test statistic. Evaluation of this conjecture would require extensive empirical studies and is not done here.

We did, however, use Monte Carlo methods to study $T(n;\theta_1,\theta_2)$ and $T(n;\hat{\theta}_1,\hat{\theta}_2)$ under the null hypothesis. The purpose was to determine the extent to which the distributions of whose statistics resembled normal distributions for small to moderate n and whether or not (a) the choice of L_n or (b) the use of parameter estimates affected the degree of approximation to the asymptotic distribution theory. The methods and results are presented in Appendix A. We concluded from the study that the asymptotic distribution theory is not adequate for determining the critical regions of the test (2.31) for moderate n (at least $n \leq 300$ and

probably much larger). Deviations from normality were particularly evident in the tails of the empirical distributions of $T(n;\theta_1,\theta_2)$ and $T(n;\hat{\theta}_1,\hat{\theta}_2)$, although those deviations may have been amplified by the limited number of replications in our study. Not surprisingly, the distributions of the statistics exhibited greater non-normal behavior when the parameters θ_1 and θ_2 were estimated from the sample than when they were known and specified. The effect of the choice of L_n on the distributions or rate of their approach to normality was not clear.

2.5 Examples

2.5.1 Obtaining the "Standard Representative" of a given Scale-Location Parameter Family

Throughout this chapter we have assumed that the distribution $G(x)$ satisfied $G^{-1}(1/2) = 0$ and $G^{-1}(3/4) - G^{-1}(1/4) = 1$. As noted previously, this assumption allowed us to use the same estimators $\hat{\theta}_1(n)$ and $\hat{\theta}_2(n)$ without our having to carry along in our calculations certain constants depending upon the particular $G(x)$ appearing in H_0. The disadvantage of this representation is that the parameters θ_1 and θ_2 are not the "usual" location and scale parameters. In this subsection we derive the equations which must be solved to determine θ_1,θ_2 and $G(x)$ given the more common form of the distribution and solve them for many useful cases.

Let $F(x) = G((x-\theta_1/\theta_2)$ be given. The relation $G^{-1}(1/2) = 0$ implies that $1/2 = G(0) = F(\theta_1)$. Now, let $x_1 = G^{-1}(1/4)$ and $x_2 = G^{-1}(3/4)$. Then, $1/4 = G(x_1) = F(\theta_2 x_1+\theta_1)$, and $3/4 = G(x_2) = F(\theta_2 x_2+\theta_1)$. Since $x_2 - x_1 = 1$, we have the following system of equations which can be solved for θ_1,θ_2 and x_1:

$$
\begin{aligned}
F(\theta_1) &= 1/2 \\
F(\theta_2 x_1+\theta_1) &= 1/4 \quad . \qquad\qquad (2.68) \\
F(\theta_2(1+x_1)+\theta_1) &= 3/4
\end{aligned}
$$

Once θ_1 and θ_2 are found, $G(x) = F(\theta_2 x+\theta_1)$ is determined. For example, suppose we are interested in the normal family $F(x) = \Phi((x-\mu)/\sigma)$. Then the equations (2.68) become

$$
\begin{aligned}
\Phi\left(\frac{\theta_1-\mu}{\sigma}\right) &= 1/2 \\
\Phi\left(\frac{\theta_2 x_1+\theta_1-\mu}{\sigma}\right) &= 1/4 \quad . \\
\Phi\left(\frac{\theta_2(1+x_1)+\theta_1-\mu}{\sigma}\right) &= 3/4
\end{aligned}
$$

The first equation above yields $\theta_1 = \mu$. The second equation then gives $\theta_2 x_1/\sigma = -.6745$; the third equation gives $\theta_2(1+x_1)/\sigma = .6745$, and together they imply $\theta_2 = 1.349$. Thence, $G(x) = \Phi(1.349x)$ and $g(x) = 1.349\phi(1.349x)$. Table 2.2 lists the results of the above transformation applied to five common distributions.

2.5.2 Weibull Alternatives

A common way to study the asymptotic power of a test is to choose a parametric family of alternatives and examine the behavior of the test as one or more of the parameters approach limiting values. Since we are dealing with a composite hypothesis which includes an entire scale-location parameter family, we must go to a three parameter family in order to study

TABLE 2.2

STANDARD REPRESENTATIVES OF SCALE-LOCATION PARAMETER FAMILIES

Distribution	$F(x) = G((x-\theta_1/\theta_2)$	$G(x)$	θ_1	θ_2	B_1	B_2
Normal	$\Phi(\frac{x-\mu}{\sigma})$ $(-\infty < x < \infty)$	$\phi(1.349x)$	μ	1.349σ	$-\infty$	$+\infty$
Exponential	$1 - \exp\{(x-\alpha)/\beta\}$ $(\alpha < x < \infty)$	$1 - (1/2)(1/3)^x$	$-\beta\ln 2 + \alpha$	$\beta\ln 3$	$-\frac{\ln 2}{\ln 3}$	$+\infty$
Extreme Value	$1 - \exp\{-\exp\{(x-\alpha)/\beta\}\}$ $(-\infty < x < \infty)$	$1 - (1/2)^{[\frac{\ln 4}{\ln(4/3)}]^x}$	$\beta\ln\ln 2 + \alpha$	$\beta\{\ln\ln 4 - \ln\ln(4/3)\}$	$-\infty$	$+\infty$
Cauchy	$\frac{1}{\pi}[\arctan(\frac{x-\alpha}{\beta}) + \frac{\pi}{2}]$ $(-\infty < x < \infty)$	$\frac{1}{\pi}[\arctan(2x) + \frac{\pi}{2}]$	α	2β	$-\infty$	$+\infty$
Logistic	$[1 + \exp\{-(x-\alpha)/\beta\}]^{-1}$ $(-\infty < x < \infty)$	$[1 + (1/9)^x]^{-1}$	α	$\beta\ln 9$	$-\infty$	$+\infty$

the asymptotic power of the test (2.31) using this approach. A natural family is the Weibull family, which approaches the exponential family as the "shape" parameter converges to one. This family has proved to be particularly well suited to the metric arising in tests based on sample spacings (Blumenthal [2]); however, the computations for the test (2.31) are considerably more tedious.

We define the sequence of alternatives

$$H_n(x) = \begin{cases} 1 - e^{-x^{\beta_n}} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where β_n is the "shape" parameter, and where, without loss of generality, we have taken the location and scale parameters to be 0 and 1 respectively. Then

$$\varepsilon_n(x) = e^{-x} - e^{-x^{\beta_n}} \qquad (x > 0)$$

and

$$h_n(x) = \beta_n x^{\beta_n - 1} e^{-x^{\beta n}} \qquad (x > 0).$$

We first consider the assumption (2.40). Defining $a(n)$ as $L_n^{1/2} K_n^{1/4}$, we will ow that

$$\beta_n = 1 - \frac{r}{a(n)} \text{ for some } r,\ 0 < r < \infty$$

satisfies the assumption, with the value of r to be determined below. Consider the function

$$S(x,n) \equiv \frac{[\varepsilon_n'(x)]^2}{h_n(x)} = h_n(x) - 2e^{-x} + [1 - \frac{r}{a(n)}]x^{-(r/a(n))}e^{(x^{\beta_n}-x)}e^{-x}.$$

Since

$$\int_0^\infty x^{-(r/a(n))}e^{-x}dx = \Gamma(1 - \frac{r}{a(n)})$$

is finite for $r < a(n)$, $S(x,n)$ is clearly integrable for all r such that $a(n) > r$. Furthermore, $S(x,n)$ is easily shown to be strictly decreasing in n, and we show below that

$$\lim_{n\to\infty} a^2(n)S(x,n) = 2r^2(x-1)^2(\log x)^2e^{-x} \tag{2.69}$$

with (see [14])

$$\int_0^\infty (x-1)^2(\log x)^2e^{-x}dx = \gamma^2 - 2\gamma + 2 + \frac{\pi^2}{6}$$

($\gamma = .5772157$ is Euler's constant). Hence, we may use the dominated convergence theorem to write

$$\lim_{n\to\infty} a^2(n)\int_0^\infty S(x,n)dx = \int_0^\infty \lim_{n\to\infty} a^2(n)S(x,n)dx = (5.6474)r^2$$

and choosing $r = .4208d_3$ satisfies (2.40).

To prove (2.69), we note that for fixed x we have

$$e^{-x^{\beta_n}} = e^{-[x - \frac{r}{a(n)}x\log x + o(\frac{r}{a(n)})]}$$

$$= e^{-x}e^{(r/a(n))x\log x\,(1+o(1))}$$

as $n \to \infty$. Hence, neglecting smaller order terms,

$$a^2(n)S(x,n) = a^2(n)e^{-x}[x^{(r/a(n))(x-1)} + x^{(r/a(n))(1-x)} - 2].$$

In order to determine the limit of the above, we treat n as a continuous variable and denote $\frac{da(n)}{dn}$ by $a'(n)$. Using the fact that

$$\frac{d}{dn}x^{(r/a(n))(1-x)} = -\frac{r}{a^2(n)}a'(n)(x-1)(\log x)x^{(r/a(n))(1-x)},$$

we apply L'Hopital's rule twice and obtain

$$\lim_{n\to\infty} a^2(n)S(x,n) = \lim_{n\to\infty}\frac{S(x,n)}{a^{-2}(n)}$$

$$= \lim_{n\to\infty}\frac{re^{-x}(\log x)[(x-1)x^{(r/a(n))(x-1)} + (1-x)x^{(r/a(n))(1-x)}]}{a^{-1}(n)}$$

$$= \lim_{n\to\infty} r^2e^{-x}(\log x)^2[(x-1)^2x^{(r/a(n))(x-1)} + (1-x)^2x^{(r/a(n))(1-x)}]$$

$$= 2r^2(x-1)^2(\log x)^2e^{-x}$$

as desired.

For this example, however, it is assumption (2.39), not (2.40), which actually determines the power of the test. In particular, we will now show that (2.39) requires

$$\beta_n = 1 - \frac{\bar{r}}{\bar{a}(n)}$$

where $\bar{a}(n) = a(n)/p_n^{1/2}$ and $\bar{r}$ is positive and finite.

Using the "standard representative" $G(x) = 1 - (1/2(1/3)^x$ from Table 2.2, tedious calculations yield the quantities appearing in definition (2.38). These expressions, neglecting smaller order terms, are summarized in Table 2.3.

TABLE 2.3

EXPRESSIONS REQUIRED FOR WEIBULL POWER COMPUTATION

Expression	Value	Expression	Value
$g(y)$	$\frac{1}{2}(\log 3)(\frac{1}{3})^y$	$H_n^{-1}(x)$	$(-\log(1-x))^{1/\beta_n}$
$G^{-1}(x)$	$-\frac{\log (2(1-x))}{\log 3}$	$\varepsilon_n(H_n^{-1}(x))$	$\frac{\bar{r}}{\bar{a}(n)}(\log \frac{1}{1-x})^2 \log\log \frac{1}{1-x}$
$g(G^{-1}(x))$	$(1-x)\log 3$	$\nu_1(n)$	$2\frac{\bar{r}}{\bar{a}(n)} \frac{(\log 2)^2 \log\log 2}{\log 3}$
$g(G^{-1}(x))G^{-1}(x)$	$(x-1)\log(2(1-x))$	$\nu_2(n)$	$4\frac{\bar{r}}{\bar{a}(n)}\left\{\frac{(\log 4)^2 \log\log 4}{\log 3} - \frac{(\log 4/3^2 \log\log 4/3}{3\log 3}\right.$
$\omega(n)$	$\log 2$		

Using the above expressions, (2.39) becomes

$$d_1 = 2\bar{r}\,\frac{(\log 2)^3 \log\log 2}{\log 3}$$

$$d_2 = 4\bar{r}\left[\frac{(\log 4)^2 \log\log 4}{\log 3} - \frac{(\log 4/3)^2 \log\log 4/3}{3\log 3}\right] \log 2.$$

Since $\nu_1^* = \lim_{n\to\infty} a(n)\nu_1(n)$ and $\nu_2^* = \lim_{n\to\infty} a(n)\nu_2(n)$ are zero, it is easily seen that the only contribution to the quantity γ_g comes from expression (2.61). Specifically,

$$\gamma_g = \lim_{n\to\infty} a^2(n)\left\{\begin{array}{l} \{\frac{(1-p_n)\log 3}{\bar{a}(n)\log 2}[d_1 - d_2\frac{\log(2(1-p_n))}{\log 3}]^2 \\ + \{\frac{p_n \log 3}{\bar{a}(n)\log 2}[d_1 - d_2\frac{\log(2p_n)}{\log 3}]\}^2 \end{array}\right\}$$

$$= [\frac{\log 3}{\log 2} d_1 - d_2]^2$$

$$= 3.4212\ \bar{r}^2,$$

and the asymptotic power of the test (2.31) is given by

$$\Phi(c_{\Phi,\alpha} + 2.419\ \bar{r}^2).$$

2.5.3 Normal Contamination Alternatives

We consider the sequence of alternatives

$$H_n(x) = (1-\delta_n)\Phi(x) + \delta_n\Phi(x+p_n)$$

where the sequence $\{\delta_n\}$ converges to zero as $n \to \infty$. These alternatives provide a basis of comparison of the test (2.31) with tests designed specifically for the normal hypothesis which we discuss in the next chapter. Using Taylor series expansions (see example 3.1), we have

$$\varepsilon_n(x) = \delta_n\{\Phi(x+p_n) - \Phi(x)\}$$

$$= \delta_n p_n\{1 - (P_n/2)x + (p_n^2/6)(x^2-1)\}\Phi(x) + o(\delta_n p_n^3)$$

and

$$\varepsilon_n'(x) = \delta_n p_n\{-x + (p_n/2)(x^2-1)\}\phi(x) + o(\delta_n p_n^2).$$

It is easily shown that condition (2.40) requires $\delta_n = d_3^{1/2}L_n^{-1/2}K_n^{-1/4}p_n^{-1}$, condition (2.39) is satisfied with $d_1 = d_2 = 0$, and (2.41) is satisfied for $b(n) = (nL_n p_n)^{1/4}$. Considering the quantities required for the power computation, $\nu_1(n) = \delta_n p_n(1+o(1))$ and $\nu_2 = o(\delta_n p_n)$; so that $\nu_1^* = d_3^{1/2}$ and $\nu_2^* = 0$. Thus, in addition to (2.61), (2.64) and (2.65) make no contribution to γ_g for this example. Also, using (2.67) and taking $G(x)$ from Table 2.2,

$$\int_{p_n}^{1-p_n} \frac{\varepsilon_n'(H_n^{-1}(x))}{h_n(H_n^{-1}(x))} \frac{g'(G^{-1}(x))}{g(G^{-1}(x))}dx = (1.349)^2\delta_n p_n \int_{p_n}^{1-p_n} [1 - (P_n/2)\Phi^{-1}(x)]\Phi^{-1}(x)dx$$

$$+ o(\delta_n p_n^2)$$

$$= (1/2)(1.349)^2\delta_n p_n^2 + o(\delta_n p_n^2)$$

from which it follows that (2.63) vanishes. Thus

$$\gamma_g = d_3 + (\nu_1^*)^2 \int_{-\infty}^{\infty} (1.349)^5 x^2\phi(1.349x)dx$$

$$= d_3(1 + (1.349)^2)$$

and the asymptotic power of the test (2.31) is given by

$$\Phi(c_{\Phi,\alpha} + 1.99d_3).$$

CHAPTER III

LARGE-SAMPLE TESTS FOR NORMALITY

3.0 Introduction

In both theoretical and applied statistics, the most common distributional assumption which is made is that of normality, and a great deal of research has been directed toward testing the validity of that assumption. In this chapter, we develop a test based on a gradually increasing number of order statistics which is a large-sample analog of the Wilk-Shapiro test for normality.

In section 3.1, we survey the literature concerning the Wilk-Shapiro and related tests.

In section 3.2, we describe the statistic $\tilde{W}(n)$. Using the asymptotic distribution theory discussed in chapter 2, we determine the asymptotic distribution of $\tilde{W}(n)$ under the null hypothesis and use that distribution to define our test for normality.

In section 3.3, we prove that the $\tilde{W}$-test is consistent.

In section 3.4, we determine the manner and rate of approach to the null hypothesis that a sequence of contiguous alternatives must satisfy in order to yield a non-trivial power as the sample size n tends to infinity.

In section 3.5, we use the results of the previous sections to define and analyze three additional statistics which give rise to improved tests for normality.

In section 3.6, we give examples of the behavior of these tests under specific sequences of alternatives.

In section 3.7, we discuss the foregoing results as well as possible

CHAPTER III

LARGE-SAMPLE TESTS FOR NORMALITY

3.0 Introduction

In both theoretical and applied statistics, the most common distributional assumption which is made is that of normality, and a great deal of research has been directed toward testing the validity of that assumption. In this chapter, we develop a test based on a gradually increasing number of order statistics which is a large-sample analog of the Wilk-Shapiro test for normality.

In section 3.1, we survey the literature concerning the Wilk-Shapiro and related tests.

In section 3.2, we describe the statistic $\tilde{W}(n)$. Using the asymptotic distribution theory discussed in chapter 2, we determine the asymptotic distribution of $\tilde{W}(n)$ under the null hypothesis and use that distribution to define our test for normality.

In section 3.3, we prove that the $\tilde{W}$-test is consistent.

In section 3.4, we determine the manner and rate of approach to the null hypothesis that a sequence of contiguous alternatives must satisfy in order to yield a non-trivial power as the sample size n tends to infinity.

In section 3.5, we use the results of the previous sections to define and analyze three additional statistics which give rise to improved tests for normality.

In section 3.6, we give examples of the behavior of these tests under specific sequences of alternati"" .

In section 3.7, we discuss the foregoing results as well as possible

areas of future research.

3.1 Goodness of Fit Tests for Normality

Since the appearance of the Wilk-Shapiro test eleven years ago, considerable attention has been devoted to testing the compound hypothesis that a sample of data comes from a population whose underlying c.d.f. is an unspecified member of the normal family. The primary focus of such research has been to understand and/or improve upon the original W-test. We begin by discussing this test, follow that with a discussion of several of the related tests which were subsequently proposed, and throughout, give a brief survey of the empirical studies which have provided the fuel for research in this area.

3.1.1 The Wilk-Shapiro W-test

The Wilk-Shapiro test was developed as a means of formalizing a rather informal procedure for testing fit--the probability plot. Briefly, a probability plot is a regression of the ordered observations on the expected values of the order statistics from a standardized version of the hypothesized distribution.

Suppose $X_1, X_2, \ldots, X_n$ is a random sample with common distribution function $G(x)$ and denote the ordered sample values by $Y_1 \leqq Y_2 \leqq \cdots \leqq Y_n$. Also suppose that the null hypothesis is true, i.e. $G(x) = \Phi(\frac{x-\mu}{\sigma})$. Then we can write

$$Y_i = \mu + \sigma Z_i \qquad i = 1,2,\ldots,n$$

where $Z_1, Z_2, \ldots, Z_n$ are the order statistics of a sample from a standard

normal population. Shapiro and Wilk [27] used the square of the generalized least squares estimator of σ obtained from the above linear model (i.e. the square of the slope of the regression line), divided that by the usual symmetric sum of squares about the mean, standardized the resulting statistic, and obtained

$$W = \frac{\dfrac{(m'V^{-1}Y)^2}{m'V^{-1}V^{-1}m}}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}$$

where

$$m_i = E[Z_i] \qquad V_{ij} = \mathrm{cov}(Z_i, Z_j)$$

$$m' = (m_1, \ldots, m_n) \qquad V = [V_{ij}].$$

Some elementary facts about W are that it lies between 0 and 1 and is location and scale invariant. Under the null hypothesis, the numerator and denominator of the above statistic are, up to a constant, estimating the same quantity, σ^2. The distribution of W under the null hypothesis was computed empirically using Monte Carlo methods and tabulated for $n = 3(1)50$. Under the alternative hypothesis, Shapiro and Wilk found, again empirically, that the values of the statistic tended to be smaller than under the null hypothesis. No analytical results have been obtained, either under H_0 or H_1, for finite n or asymptotically as n tends to infinity. But in empirical studies, the W-test, which rejects H_0 for small values of W, has been shown to be at least as good and generally better than most of the commonly used tests over various symmetric,

asymmetric, short- and long-tailed alternatives.

By far the most extensive of these empirical power studies was reported by Shapiro, Wilk, and Chen [29]. In that study, they compared W with eight other tests, including the Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling, over twelve families of alternative distributions. They found the W-test to be a more sensitive indicator of non-normality than any of the other tests; in addition to exhibiting better power against most alternatives, it was the only test which never had a low power when another test had a high power against a continuous alternative distribution. Stephens [30] later amended their results in comparing W to the various EDF tests mentioned above. He pointed out that the study [29] computed the EDF statistics by treating the means and variances of the alternative distributions as if they were completely specified rather than using estimates obtained from the samples. Stephens still found, however, that the W-test was generally superior to the best EDF competitor, though usually only slightly. Chen [3] studied the behavior of W under scale and location contaminated normal distributions, finding the W-test to be sensitive to both types of deviations from normality, although the power dropped off rapidly as the contaminating distributions approached the normal.

The primary drawbacks to the use of the W-test are: 1) the coefficients cannot be computed, but must be obtained from tables, a different set of coefficients for each different value of n; and 2) the coefficients (in particular the v_{ij}'s) are only known exactly for $n \leq 20$. (Shapiro and Wilk used linear regression to extrapolate the coefficient values up to $n = 50$.)

In a later paper [28], Shapiro and Wilk found good empirical agreement between the distribution of W and a member of Johnson's S_B class of distributions.

3.1.2 Related Tests

The nonavailability of the v_{ij}'s for $n \geqq 20$ prompted Shapiro and Francia [26] to propose the following statistic as an approximation to W:

$$W^* = \frac{\dfrac{(m'Y)^2}{m'm}}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2} .$$

They justified the use of W^* on the grounds that, for large n, the order statistics $\{Y_i\}$ could be treated as if they were independent. This, of course, is false, but Stephens [31] later showed that in the normal case the vector of mean values of the standardized order statistics is an asymptotic eigenvector of the covariance matrix of those statistics. Thus, the statistic W^* is in fact asymptotically related to W, and has the advantage that the vector m is known up to $n = 400$. Shapiro and Francia have computed the critical region for the W^*-test, which rejects for small values of W^*, for $n = 35,50(1)99$. Again, no analytical results are available, but in empirical studies [26,30,35], the W^*-test has been shown to be very similar, but, in general, slightly inferior in performance to the W-test, although Shapiro and Francia found the W^*-test to be more sensitive to "near normal" alternatives. Sarkadi [22] proved the consistency of the W^*-test against alternatives with finite second moments.

The most extensive analytical work in this area has been done by De Wet and Venter [9,10], who were concerned with the statistic

$$r_n^2 = \frac{\dfrac{(\overline{m}'Y)^2}{\overline{m}'\overline{m}}}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}$$

where $\overline{m}' = (\overline{m}_1,\ldots,\overline{m}_n)$ with $\overline{m}_i = \Phi^{-1}(\frac{i}{n+1})$. They determined the asymptotic distribution of $n(1 - r_n^2) - a_n$ under the null hypothesis to be $\Psi(y)$, where $\Psi(y)$ is the distribution of

$$\sum_{k=3}^{\infty} \frac{(X_k^2 - 1)}{k}$$

where $X_3, X_4, \ldots$ are i.i.d. standard normal random variables, and where a_n is a complicated function which is $o(\log n)$ as n tends to infinity. Their methods are discussed in more detail in Appendix B. The percentiles of $\Psi(y)$ were computed numerically using the inverse of the characteristic function. As we might expect since the $\overline{m}_i$ are limiting values of the m_i, good agreement was found between r_n^2 and W^* (for $n = 99$); it would b· difficult, however, to make a rigorous statement concerning the re. onship between the two statistics. Further remarks about r_n^2 and W^* as well as additional empirical approximations to the distributions of W and W^* appear in Stephens [33].

In other work building on his earlier results on covariance matrices, Stephens [32] considered the more general linear model

$$E(Y_i) = \mu + \sigma m_i + \beta_2 w_2(m_i) + \beta_3 w_3(m_i) + \ldots$$

where $\beta_2, \beta_3, \ldots$ are constants and $w_2(m_i), w_3(m_i), \ldots$ are functions of the m_i, in particular, the Hermite polynomials. Under the null hypothesis, $0 = \beta_2 = \beta_3 = \ldots$, and a large-sample test of the latter event is proposed which may be extended to as many β_i as the experimenter wishes. Specifically, by choosing $w_2(.), w_3(.), \ldots$ as he did, Stephens was able to obtain simple expressions for the estimates $\hat{\beta}_2, \hat{\beta}_3, \ldots$ and, under the null hypothesis, show that standardized, studentized versions of these estimates are asymptotically independent and normally distributed. Thus, the sum of the square estimates, $\hat{\beta}_2^2 + \hat{\beta}_3^2 + \ldots$ has a chi-square distribution with the number of degrees of freedom determined by the number of estimators which the experimentor wishes to consider. Stephens' test is based on the latter distribution.

Finally, D'Agostino [5] considered tests based on the following statistic:

$$D_A = \frac{\sum_{i=1}^{n} [i - (1/2)(n+1)] Y_i}{[n^3 \sum_{i=1}^{n} (Y_i - \overline{Y})^2]^{1/2}} .$$

D'Agostino showed that under the null hypothesis, D_A is asymptotically normally distributed, but he found that n must be well over 1000 before the asymptotic distribution is useful in determining the critical regions of his test. Instead, he used Cornish-Fisher expansions (based on the first four cumulants of D_A) to determine the percentiles of the distribution of a standardized version of D_A. In an empirical power study, he found that a two-sided test based on D_A was able to detect deviations from normality due to both skewness and kurtosis (i.e. the test is an

"omnibus" test). He also found that this test was comparable, but again slightly inferior, to the W-test. In a later study, Theune [35] found the Wilk-Shapiro test to be "far superior" to D'Agostino's test in detecting non-normality. Theune did, however, find the D'Agostino test to be comparable to the Shapiro-Francia test over most of the alternative distributions studied.

3.2 A Large-Sample Test for Normality Related to the W-Test

In this section we describe a statistic $\tilde{W}(n)$ which is an analog of the Shapiro-Wilk W based on an increasing subset of the order statistics. By appropriately selecting our vector of order statistics as in the previous chapter, we are able to invoke the asymptotic distribution theory of section 2.2.4 and thereby determine the asymptotic behavior of $\tilde{W}(n)$. Based on the asymptotic distribution of $\tilde{W}(n)$ under the null hypothesis, we will define the α-level critical region for our test. In view of the asymptotic sufficiency result of Weiss [40], we would hope that, at least for large n, the powers of an α-level test based on $\tilde{W}(n)$ and an α-level test based on W (assuming the necessary coefficients were available) would be "close" in some sense. In that event, the large sample behavior of the $\tilde{W}$-test would help to explain the behavior of the W-test, and at the same time, the empirically determined superiority of the W-test for small samples would provide some justification for using the $\tilde{W}$-test when n is large. The relationship between the two tests remains open question. Some further remarks concerning this relationship appear in subsection 3.2.4.

Before defining our test statistic, we recall some notation and results from chapter 2.

3.2.1 Definitions

Let the sequences of integers $\{np_n\}$, $\{K_n\}$, and $\{L_n\}$ be as defined in section 2.2.1 with the assumptions (2) through (8) replaced by

$$\lim_{n\to\infty} p_n = 0 \qquad \lim_{n\to\infty} K_n = \infty \qquad \lim_{n\to\infty} np_n = \infty \tag{3.1}$$

$$\lim_{n\to\infty} \frac{L_n}{np_n^2 \log(1/p_n)} = 0 \tag{3.2}$$

$$\lim_{n\to\infty} \frac{n}{L_n^{3/2} p_n^2 \log(1/p_n)} = 0 \tag{3.3}$$

$$\lim_{n\to\infty} \frac{n^3 p_n}{L_n^4} = 0 \tag{3.4}$$

We note that (3.2) and (3.3) above are exactly (2.4) - (2.6) where we have replaced τ_n by its value for large n (up to a constant) in the normal case as given by (C.10). The assumption (3.4) is new. For $L_n \geqq n^{3/4}$ it clearly imposes no added restrictions on our choice of parameters. For $L_n = o(n^{3/4})$, it requires in effect that our selected subset of order statistics move into the tails fairly rapidly as n gets large.

Throughout this chapter we will denote

$$\Phi^{-1}\left(\frac{np_n+(i-1)L_n}{n}\right)$$

by $T_i(n)$ and $\phi(T_i(n))$ by $t_i(n)$, for $i = 1,2,\ldots,K_n+1$. Also, we will denote $\Phi^{-1}(x)$ by $T(x)$ and $\phi(T(x))$ by $t(x)$. We will denote the (K_n+1)-vector of selected order statistics by $\tilde{Y}(n)$, with components given by

$$\tilde{Y}_i \equiv \tilde{Y}_i(n) = \begin{cases} q_n Y_{np_n} & i=1 \\ Y_{np_n+(i-1)L_n} & i=2,3,\ldots,K_n \\ q_n Y_{n(1-p_n)} & i=K_n+1 \end{cases}$$

where

$$q_n \equiv \{\frac{n}{2L_n}\frac{t_1(n)}{T_1(n)}[\frac{t_1(n)T_1(n)}{p_n} + T_1^2(n) - 1 + \frac{L_nT_1(n)}{n\;t_1(n)} + \frac{1}{3}(\frac{L_n}{n})^2(\frac{1+T_1^2(n)}{t_1^2(n)})]\}^{1/2}. \tag{3.5}$$

Let $\tilde{X}(n) = (\tilde{X}_1,\ldots,\tilde{X}_{K_n+1})$ be the analogous order statistics selected from a sample whose parent population is standard normal. Then, in view of the asymptotic distribution theory of section 2.1.4, we can assume for all asymptotic probability calculations that the vector $\tilde{X}(n)$ has a (K_n+1)-variate normal distribution with mean vector

$$\tilde{T}(n) = (\tilde{T}_1(n),\ldots,\tilde{T}_{K_n+1}(n))' \equiv (q_nT_1(n),T_2(n),\ldots,T_{K_n}(n),q_nT_{K_n+1}(n))'$$

and covariance matrix $V(n)$ whose components $v_{ij}(n)$ are given by

$$v_{ij}(n) = \frac{L_n}{n(L_n-1)}\begin{cases} q_n^2p_n(1-p_n)/t_1^2(n) & i=j=1 \\ q_n^2p_n^2/t_1(n)t_{K_n+1}(n) & i=1,j=K_n+1 \\ q_np_n(1-p_n-(j-1)\frac{L_n}{n})/t_1(n)t_j(n) & 1=i<j\leqq K_n \\ (p_n+(i-1)\frac{L_n}{n})(1-p_n-(j-1)\frac{L_n}{n})/t_i(n)t_j(n) & 2\leqq i\leqq j\leqq K_n \\ q_n(p_n+(i-1)\frac{L_n}{n})p_n/t_i(n)t_{K_n+1}(n) & 2\leqq i<j=K_n+1 \\ q_n^2p_n(1-p_n)/t_{K_n+1}^2(n) & i=j=K_n+1 \end{cases} \tag{3.6}$$

The inverse of the above covariance matrix has the simpler form:

$$V^{-1}(n)=\frac{n^2(L_n-1)}{L_n}\begin{bmatrix} \frac{1+\frac{L_n}{np_n}}{q_n^2}t_1^2 & \frac{t_1t_2}{q_n} & & & & \\ \frac{-t_1t_2}{q_n} & 2t_2^2 & -t_2t_3 & & & \\ & -t_2t_3 & 2t_3^2 & -t_3t_4 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -t_{K_n-1}t_{K_n} & 2t_{K_n}^2 & -t_{K_n}t_{K_n+1} \\ & & & & \frac{-t_{K_n}t_{K_n+1}}{q_n} & \frac{1+\frac{L_n}{np_n}}{q_n^2}t_{K_n+1}^2 \end{bmatrix} . \quad (3.7)$$

3.2.2 The $\tilde{W}$-Test

We are concerned with the testing problem P_n of section 2.0, with $G(x) = \Phi(x)$, the standard normal c.d.f. Analogous to the Wilk-Shapiro statistic, we define our test statistic $\tilde{W}(n)$ as

$$\tilde{W}(n) = \frac{[\tilde{T}'(n)V^{-1}(n)\tilde{Y}(n)]^2}{[\tilde{T}'(n)V^{-1}(n)V^{-1}(n)\tilde{T}(n)]\tilde{D}^2(n)} \qquad (3.8)$$

where

$$\tilde{D}^2(n) = \sum_{i=1}^{K_n+1}[\tilde{Y}_i(n) - \tilde{Y}.(n)]^2 \quad \text{with} \quad \tilde{Y}.(n) = (K_n+1)^{-1}\sum_{i=1}^{K_n+1}\tilde{Y}_i(n).$$

Under H_0, the numerator of $\tilde{W}(n)$ may be thought of as the least squares

estimator of the variance σ^2 based on $\tilde{Y}(n)$. The denominator, as we shall prove later, has the same asymptotic distribution (after suitable standardization) as the denominator of W.

Lemma 3.1. For all n,

(i) $\tilde{W}(n)$ is location and scale invariant

(ii) $0 \leq \tilde{W}(n) \leq 1$.

Proof. Identical to the proofs of Lemmas 1 and 2 of Shapiro and Wilk [27].

Lemma 3.2. As $n \to \infty$,

$$q_n = \left(\frac{np_n}{L_n}\right)^{1/2} (1+o(1)) \tag{3.9}$$

Proof. The proof is based on the following formula of Feller [13,p.193]:

$$\frac{1}{x} - \frac{1}{x^3} + \frac{3}{x^5} - \frac{15}{x^7} < \frac{1-\Phi(x)}{\phi(x)} < \frac{1}{x} - \frac{1}{x^3} + \frac{3}{x^5} \tag{3.10}$$

for all $x > 0$. Taking $x = \Phi^{-1}(1-p_n) = -T_1(n)$ in (3.10), we have $\phi(x) = t_1(n)$ and (3.10) becomes

$$-\frac{1}{T_1(n)} + \frac{1}{T_1^3(n)} - \frac{3}{T_1^5(n)} + \frac{15}{T_1^7(n)} < \frac{p_n}{t_1(n)} < -\frac{1}{T_1(n)} + \frac{1}{T_1^3(n)} - \frac{3}{T_1^5(n)}. \tag{3.10'}$$

Multiplying both sides of (3.10') by $t_1(n)T_1^2(n)/p_n$ and rearranging terms yields

$$\frac{t_1(n)}{p_nT_1(n)} - \frac{3t_1(n)}{p_nT_1^3(n)} + \frac{15t_1(n)}{p_nT_1^5(n)} < \frac{t_1(n)T_1(n)}{p_n} + T_1^2(n) < \frac{t_1(n)}{p_nT_1(n)} - \frac{3t_1(n)}{p_nT_1^3(n)}. \tag{3.11}$$

Multiplying both sides of (3.10') by $t_1(n)/p_n$ and rearranging terms yields

$$\frac{t_1(n)}{p_nT_1^3(n)} - \frac{3t_1(n)}{p_nT_1^5(n)} + \frac{15t_1(n)}{p_nT_1^7(n)} < \frac{t_1(n)}{p_nT_1(n)} + 1 < \frac{t_1(n)}{p_nT_1^3(n)} - \frac{3t_1(n)}{p_nT_1^5(n)} . \tag{3.12}$$

Substituting (3.12) into (3.11) we obtain

$$\begin{aligned} -2 - \frac{2t_1(n)}{p_nT_1^3(n)} + \frac{12t_1(n)}{p_nT_1^5(n)} + \frac{15t_1(n)}{p_nT_1^7(n)} &< \frac{t_1(n)T_1(n)}{p_n} + T_1^2(n) - 1 \\ &< -2 - \frac{2t_1(n)}{p_nT_1^3(n)} - \frac{3t_1(n)}{p_nT_1^5(n)} . \end{aligned} \tag{3.13}$$

Thus, using (C.7) and (C.8),

$$\frac{t_1(n)T_1(n)}{p_n} + T_1^2(n) - 1 = -2(1 + 0(\frac{1}{2\log(1/p_n)}))$$

as $n \to \infty$. Finally, combining the above relation with the definition of q_n^2 and using (3.2) to e ιinate smaller order terms, we obtain

$$q_n^2 = \frac{np_n}{L_n}(1 + 0(\frac{1}{2\log(1/p_n)})) .$$

This completes the proof of lemma 3.2. QED

Since we are concerned here only with asymptotic results, one might ask why, in view of lemma 3.2, the definition (3.ι) was given for q_n instead of the much simpler and more easily computable $q_n^2 = np_n/L_n$? The answer is that even for large n, the smaller order terms in (3.5) play a

crucial role in the asymptotic distribution of $\tilde{W}(n)$. The exact difficulties which arise if those terms are neglected will be pointed out later. In many of the arguments below, however, it will be true that those terms may be neglected for asymptotic purposes, in which case the result of lemma 3.2 will be useful.

Let $\tilde{\mu}(n)$ denote the quantity

$$\frac{[\tilde{T}'(n)V^{-1}(n)\tilde{T}(n)]^2}{[\tilde{T}'(n)V^{-1}(n)V^{-1}(n)\tilde{T}(n)]\tilde{T}'(n)\tilde{T}(n)} .$$

The next subsection of this chapter is devoted to proving the following theorem:

<u>Theorem 3.1</u>. <u>Under the null hypothesis</u>

$$\left(\frac{27}{2}\right)^{1/2} \frac{n^{5/2}p_n^{3/2}}{L_n^2} \cdot (\tilde{W}(n) - \tilde{\mu}(n)) \tag{3.14}$$

<u>is asymptotically normally distributed with mean 0 and variance 1.</u>

Thus, we define the α-level $\tilde{W}$-test for normality as follows:

$$\text{Reject } H_0 \text{ iff } \tilde{W}(n) < \tilde{\mu}(n) + c_{\Phi,\alpha}\tilde{\sigma}(n)$$

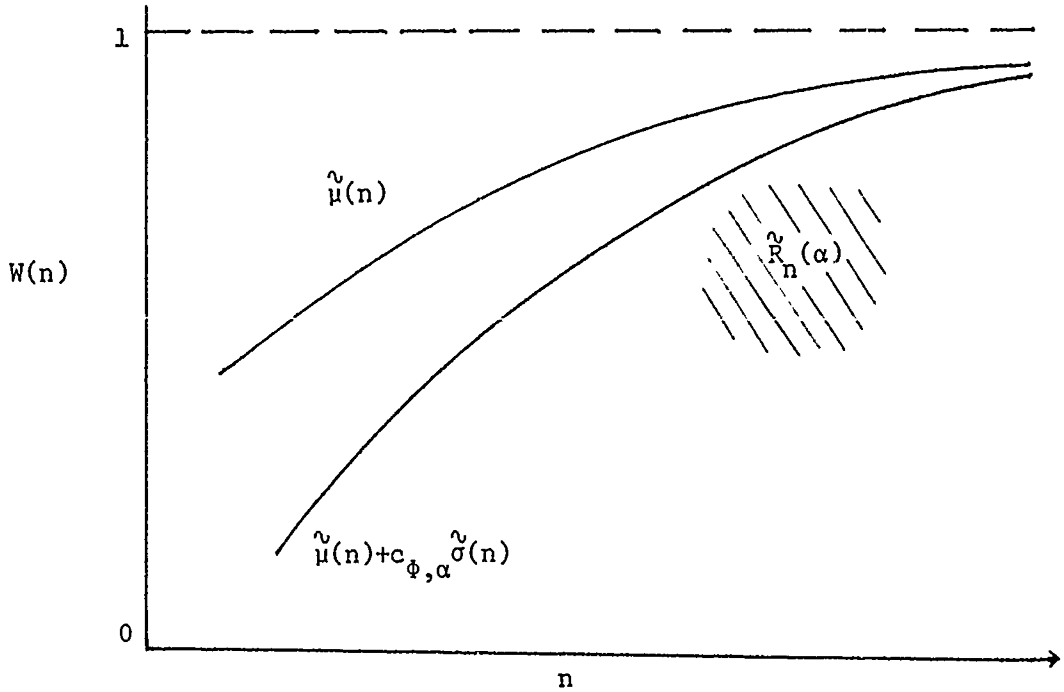
where

$$\tilde{\sigma}(n) \equiv \left(\frac{2}{27}\right)^{1/2} \frac{L_n^2}{n^{5/2}p_n^{3/2}}$$

and $c_{\Phi,\alpha}$ is the α-percentage point of the standard normal c.d.f. The reason for choosing the lower tail of the distribution as our critical region is that under the alternative hypothesis, $\tilde{W}(n)$ tends to be smaller than under the null hypothesis (this statement is made rigorous in section

FIGURE 3.1

CRITICAL REGION $\tilde{R}_n(\alpha)$ OF THE $\tilde{W}$-TEST (Schematic)



3.3). Figure 3.1 depicts the large sample behavior of $\tilde{W}(n)$ and the critical region $\tilde{R}_n(\alpha)$ of the test. In lemma 3.5 below, we show that $1 - \tilde{\mu}(n)$ is

$$O\left(\frac{L_n^4 \log(1/p_n)}{n^4 p_n^3}\right) .$$

Hence, (3.4) implies that the quantity $\tilde{\sigma}(n)$ converges to zero faster than the quantity $\tilde{\mu}(n)$ converges to 1, and the $100(1-\alpha)\%$ confidence region for $\tilde{W}(n)$ approaches the curve $\tilde{\mu}(n)$ faster than the curve approaches its limiting value.

3.2.3 Proof of Theorem 3.1

Since we are assuming that the null hypothesis holds, we may also assume without loss of generality by lemma 3.1 that $Y_1,\ldots,Y_n$ are the order statistics of a sample from a $N(0,1)$ population. Hence, for all our asymptotic probability calculations, we will hereafter take the distribution of $\tilde{Y}(n)$ to be (K_n+1)-variate normal with mean vector $\tilde{T}(n)$ and covariance matrix $V(n)$ as defined in the previous section. The theorem will be proved by examining the Taylor series expansion of $\tilde{W}(n)$ about the point $\tilde{T}(n)$ and determining which terms are asymptotically dominant. Before doing so, we introduce two lemmas which are required for the proof.

Lemma 3.3. As $n \to \infty$

$$\operatorname{var} \tilde{Y}_i = \begin{cases} \dfrac{1}{L_n \log(1/p_n)}(1+o(1)) & i=1,K_n+1 \\ \dfrac{1}{np_n \log(1/p_n)}\,O(1) & i=2,3,\ldots,K_n \end{cases} \tag{3.15}$$

and

$$\sum_{j=1}^{K_n+1} v_{ij}(n) = \begin{cases} \left(\dfrac{np_n}{L_n^3}\right)^{1/2} O(1) & i=1,K_n+1 \\ \dfrac{1}{L_n-1}(1+o(1)) & i=2,3,\ldots,K_n \end{cases} \tag{3.16}$$

Proof. It is easily verified that $\max\limits_{2\leq i\leq K_n} \{\operatorname{var} \tilde{Y}_i\}$ occurs at $i=2$ or $i=K_n$. An application of (C.10) yields (3.15).

In proving (3.16), we first consider the case $i=1$ (or $i=K_n+1$ by symmetry). Using the definition (3.6) and simplifying, we have

$$\sum_{j=1}^{K_n+1} v_{ij}(n) = \frac{p_n^2}{(L_n-1)t_1^2(n)} + \frac{p_n^{3/2}}{L_n^{1/2}n^{1/2}t_1(n)} \sum_{j=2}^{K_n} \frac{[1-p_n-(j-1)\frac{L_n}{n}]}{t_j(n)} .$$

By (C.10), the first term is $O([L_n \log(1/p_n)]^{-1})$ as $n \to \infty$, and the second term is approximately

$$\left(\frac{p_n^3}{2L_n n}\right)^{1/2} \frac{n}{L_n t_1(n)} \int_{p_n}^{1-p_n} \frac{(1-x)}{t(x)}\, dx$$

which equals

$$\left(\frac{np_n^3}{2L_n^3}\right)^{1/2} (1-2p_n)\frac{T_1(n)}{t_1(n)} .$$

Applying (C.8), the first half of (3.16) follows. (In Chapter 3 as in Chapter 2, we will frequently have occasion to approximate summation by integration as above. We again refer the reader to Appendix C, where we formally carry out this procedure and obtain upper bounds for the resulting errors for integrals of the type appearing in this chapter.)

For $i=2,3,\ldots,K_n$ we have

$$\sum_{j=1}^{K_n+1} v_{ij}(n) = q_n\left[-\frac{p_n(1-p_n-(j-1)\frac{L_n}{n})}{t_1(n)t_i(n)} + \frac{p_n(p_n+(i-1)\frac{L_n}{n})}{t_{K_n+1}(n)t_i(n)}\right]\frac{L_n}{n(L_n-1)} +$$

$$\frac{L_n}{n(L_n-1)} \left\{ \sum_{j=2}^{i-1} \frac{[p_n+(j-1)\frac{L_n}{n}][1-p_n-(i-1)\frac{L_n}{n}]}{nt_i(n)t_j(n)} \right.$$

$$+ \sum_{j=i}^{K_n} \frac{[p_n+(i-1)\frac{L_n}{n}][1-p_n-(j-1)\frac{L_n}{n}]}{nt_i(n)t_j(n)} \}$$

$$= O(\frac{L_n p_n q_n}{n(L_n-1)t_1(n)t_i(n)}) + \frac{1}{L_n-1} \{\int_{T_1(n)}^{y} \frac{\Phi(x)[1-\Phi(y)]}{\phi(y)} dx$$

$$+ \int_{y}^{T_{K_n+1}(n)} \frac{[1-\Phi(x)]\Phi(y)}{\phi(y)} dx \quad (1+o(1))\}$$

where $y \cong T_i(n)$. The first term above is $o([np_nL_n]^{-1/2})$ as $n \to \infty$. In particular, since $(np_nL_n)^{-1/2} = L_n^{-1}(L_n/np_n)^{1/2}$, it is $o(L_n^{-1})$. Noting that

$$\int_{-a}^{-y} \Phi(x)dx = \int_{y}^{a} [1-\Phi(x)]dx,$$

the second term above can be written as

$$[(L_n-1)\phi(y)]^{-1} \{ \int_{T_1(n)}^{y} \Phi(x)dx - y\Phi(y)\}.$$

Evaluating the expression in braces, we obtain $\{p_nT_1(n) + \phi(y) - t_1(n)\}$. An application of (C.8) completes the proof of (3.16). This completes the proof of lemma 3.3. QED

It is well known that for the full set (i.e. all n) of standard normal order statistics, the row sums of the covariance matrix are equal to 1. Equation (3.16) shows that the ratio of the common row sum for our selected subset of order statistics to that of the complete set is asymptotically equal to the fraction of order statistics that we choose.

<u>Lemma 3.4. The vector $\tilde{T}(n)$ is an asymptotic eigenvector of the matrix $V(n)$ corresponding to the eigenvalue</u> $[2(L_n-1)]^{-1}$.

<u>Proof</u>. We will denote the (K_n+1)-vector $\tilde{T}'(n)V^{-1}(n)$ by $\gamma(n)$, with individual components $\gamma_i(n)$. It follows directly from the definitions that

$$\gamma_i(n) = \frac{n^2(L_n-1)}{L_n^2} \begin{cases} \dfrac{t_1(n)}{q_n}\{(1+\dfrac{L_n}{np_n}t_1(n)T_1(n) - t_2(n)T_2(n)\} & i=1 \\ t_i(n)\{2t_i(n)T_i(n)-t_{i-1}(n)T_{i-1}(n)-t_{i+1}(n)T_{i+1}(n)\} & i=2,3,\ldots,K_n \\ \dfrac{t_{K_n+1}(n)}{q_n}\{(1+\dfrac{L_n}{np_n})t_{K_n+i}(n)T_{K_n+1}(n)-t_{K_n}(n)T_{K_n}(n)\} & i=K_n+1 \end{cases}$$

Using (C.1), Taylor series expansions of $t_{i-1}(n)T_{i-1}(n)$ and $t_{i+1}(n)T_{i+1}(n)$ yield

$$\begin{aligned} t_{i-1}(n)T_{i-1}(n) = t_i(n)T_i(n) &- \frac{L_n}{n}(1 - T_i^2(n)) - (\frac{L_n}{n})^2 \frac{T_i(n)}{t_i(n)} \\ &+ \frac{1}{3}(\frac{L_n}{n})^3 (\frac{1+T_i^2(n)}{t_i^2(n)}) - \frac{1}{6}(\frac{L_n}{n})^4 (\frac{T_i(n)}{t_i(n)})^3 + \delta_1(i,n) \end{aligned} \tag{3.17}$$

$$\begin{aligned} t_{i+1}(n)T_{i+1}(n) = t_i(n)T_i(n) &+ \frac{L_n}{n}(1 - T_i^2(n)) - (\frac{L_n}{n})^2 \frac{T_i(n)}{t_i(n)} \\ &- \frac{1}{3}(\frac{L_n}{n})^3 (\frac{1+T_i^2(n)}{t_i^2(n)}) - \frac{1}{6}(\frac{L_n}{n})^4 (\frac{T_i(n)}{t_i(n)})^3 + \delta_2(i,n) \end{aligned}$$

where $\delta_1(i,n)$ and $\delta_2(i,n)$ are $o(L_n^4/n^4p_n^3)$ as $n \to \infty$. Hence,

$$\gamma_i(n) = \frac{n^2(L_n-1)}{L_n^2}\begin{cases} \frac{t_1(n)}{q_n}\{\frac{L_n}{np_n}t_1(n)T_1(n) + \frac{L_n}{n}(T_1^2(n)-1) + (\frac{L_n}{n})^2\frac{T_1(n)}{t_1(n)} \\ \quad + \frac{1}{3}(\frac{L_n}{n})^3(\frac{1+T_1^2(n)}{t_1^2(n)}) + \frac{1}{6}(\frac{L_n}{n})^4(\frac{T_1(n)}{t_1(n)})^3 - \delta_2(1,n)\} & i=1 \\ t_i(n)\{2(\frac{L_n}{n})^2\frac{T_i(n)}{t_i(n)} + \frac{1}{3}(\frac{L_n}{n})^4(\frac{T_i(n)}{t_i(n)})^3 + \delta_3(i,n)\} & i=2,\ldots,K_n \\ \frac{t_{K_n+1}(n)}{q_n}\{\frac{L_n}{np_n}t_{K_n+1}(n)T_{K_n+1}(n) - \frac{L_n}{n}(T_{K_n+1}^2(n)-1) \\ \quad + (\frac{L_n}{n})^2(\frac{T_{K_n+1}(n)}{t_{K_n+1}(n)}) - \frac{1}{3}(\frac{L_n}{n})^3(\frac{1+T_{K_n+1}^2(n)}{t_{K_n+1}^2(n)}) & i=K_n+1 \\ \quad + \frac{1}{6}(\frac{L_n}{n})^4(\frac{T_{K_n+1}(n)}{t_{K_n+1}(n)})^3 - \delta_1(K_n+1,n)\} \end{cases} \tag{3.18}$$

where $\delta_3(i,n) = -(\delta_1(i,n) + \delta_2(i,n))$. We note that $\delta_3((K_n+2)/2,n) = 0$ for all n. Let $\delta_4(n)$ denote the quantity

$$\frac{t_1(n)}{q_n}\{\frac{L_n}{np_n}t_1(n)T_1(n) + \frac{L_n}{n}(T_1^2(n)-1) + (\frac{L_n}{n})^2\frac{T_1(n)}{t_1(n)} + \frac{1}{3}(\frac{L_n}{n})^3(\frac{1+T_1^2(n)}{t_1^2(n)})\}$$
$$- 2(\frac{L_n}{n})^2\tilde{T}_1(n).$$

By the definition of q_n, $\delta_4(n) \equiv 0$. Also, using (3.9) and (C.7),

$$\gamma'(n) \equiv \tilde{T}'(n)V^{-1}(n) = 2(L_n-1)(1+O[\frac{L_n}{np_n}]^2))\tilde{T}'(n).$$

Post-multiplying both sides of the above equation by $V(n)$, rearranging terms and applying (3.2), the proof of lemma 3.4 is complete. QED

Remarks. The above result is not surprising in view of the work of Stephens [31]. Working with the complete set of normal order statistics, Stephens found that the mean vector is again an asymptotic eigenvector of the covariance matrix, but this time corresponding to the eigenvalue $\gamma = 1/2$. More generally, by solving a certain differential equation, Stephens found that the sequence $\{\lambda_j = (j+1)^{-1}\}$ is a sequence of asymptotic eigenvalues for the covariance matrix of $(Y_1,\ldots,Y_n)$, with corresponding eigenvectors $\{w_j\}$ whose components are the Hermite polynomials $He_j(x)$ evaluated at the appropriate population quantiles. Stephens' method does not yield error terms, however, In our case, since the eigenvalues are approaching zero as n increases, the asymptotic result is only meaningful provided the error terms approach zero at a faster rate. Hence the above method of proof.

An Example. Before continuing with the proof of the theorem, we give a simple application of lemmas 3.3 and 3.4, using them to show that the symmetric sum of squares estimator of the variance σ^2 and its analog based on $\tilde{Y}(n)$, suitably standardized, have the same asymptotic distribution. Without loss of generality, we take $\sigma^2 = 1$ as above. Let $S^2(n)$ denote

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2.$$

It is well know that $S^2(n)$ is approximately $N(1,2/n)$ for large n, and

thus, $(n/2)^{1/2}(S^2(n) - 1)$ converges in law to the standard normal distribution. Now consider

$$\tilde{S}^2(n) = (K_n+1)^{-1} \sum_{i=1}^{K_n+1} (\tilde{Y}_i - \tilde{Y}.)^2$$

Expanding $\tilde{S}^2(n)$ in a Taylor series about $\tilde{T}(n)$,

$$\tilde{S}^2(n) = (K_n+1)^{-1} \sum_{i=1}^{K_n+1} \tilde{T}_i^2(n) + 2(K_n+1)^{-1} \sum_{i=1}^{K_n+1} \tilde{T}_i(n)(\tilde{Y}_i - \tilde{T}_i(n))$$

$$+ (K_n+1)^{-1} \sum_{i=1}^{K_n+1} \sum_{j=1}^{K_n+1} (\delta_{ij}-(K_n+1)^{-1})(\tilde{Y}_i-\tilde{T}_i(n))(\tilde{Y}_j-\tilde{T}_j(n))$$

where $\delta_{ij} = 1$ if $i = j$ and $= 0$ otherwise.

Using (C.9), the first term in the above expression is easily shown to converge to 1 as n tends to infinity:

$$(K_n+1)^{-1} \sum_{i=1}^{K_n+1} \tilde{T}_i^2(n) = (K_n+1)^{-1}[2q_n^2 T_1^2(n) + \sum_{i=2}^{K_n} T_i^2(n)]$$

$$= 4p_n \log(1/p_n)(1+o(1)) + \int_{p_n}^{1-p_n} T^2(x)dx(1+o(1))$$

$$= 4p_n \log(1/p_n)(1+o(1)) + (1-2p_n-4p_n \log(1/p_n))(1+o(1))$$

$$= 1 - o(p_n \log(1/p_n)).$$

The second term is asymptotically normal with mean 0 and variance $4K_n^{-2}\tilde{T}'(n)V(n)\tilde{T}(n)$, which, by lemma 3.4, can be written for asymptotic purposes as $2n^{-1}$. And finally, since the third term has expectation

$o((np_n)^{-1})$ (by lemma 3.3) and standard deviation also $o((np_n)^{-1})$ (by an argument analogous to the one used in the proof of lemma 3.7 below), it is $o_p((np_n)^{-1})$ as $n \to \infty$. Thus, $(n/2)^{1/2}(\tilde{S}^2(n) - E\{S^2(n)\})$ converges in law to the standard normal distribution, where $E\{S^2(n)\}$ converges to 1 as $n \to \infty$.

Proceeding now with the Taylor series expansion of $\tilde{W}(n)$, we have

$$\tilde{W}(n) = \frac{[\tilde{T}'(n)V^{-1}(n)\tilde{T}(n)]^2}{[\tilde{T}'(n)V^{-1}(n)V^{-1}(n)\tilde{T}(n)][\tilde{T}'(n)\tilde{T}(n)]} + \sum_{i=1}^{K_n+1} \tilde{w}_i^{(1)}(\tilde{T}(n))(\tilde{Y}_i - \tilde{T}_i(n))$$
$$+ \sum_{i=1}^{K_n+1} \sum_{j=1}^{K_n+1} \tilde{w}_{ij}^{(2)}(\tilde{T}(n))(\tilde{Y}_i - \tilde{T}_i(n))(\tilde{Y}_j - \tilde{T}_j(n)) + \dots \tag{3.19}$$

where $\{\tilde{w}_i^{(1)}(\tilde{T}(n))\}$ denotes the set of first partial derivatives evaluated at $\tilde{T}(n)$, $\{\tilde{w}_{ij}^{(2)}(\tilde{T}(n))\}$ the set of second partials, etc. To simplify notation, we will hereafter denote the first (constant) term in the expansion (3.19) by $S_1(n)$, the second (linear) term by $S_2(n)$, the third (quadratic) term by $S_3(n)$, and the remaining terms $(\tilde{W}(n) - S_1(n) - S_2(n) - S_3(n))$ by $S_4(n)$. In lemmas 3.5 to 3.7, we deal with $S_1(n)$, $S_2(n)$ and $(S_3(n)+S_4(n))$ respectively.

Lemma 3.5. As $n \to \infty$,

$$S_1(n) = 1 - \frac{L_n^4}{27n^4p_n^3}\log(1/p_n)\,(1+o(1)). \tag{3.20}$$

Proof. Contained in the proof of lemma 3.4 is the fact that

$$\gamma'(n) = 2(L_n-1)T'(n) + e'(n)$$

where $e'(n) = (e_1(n),\ldots,e_{K_n+1}(n))$, with

$$e_i(n) = \begin{cases} (L_n-1)\delta_5(n)(1 + o(1)) & i=1 \\ \frac{1}{3}\,\frac{L_n^2(L_n-1)}{n^2}\,\frac{T_i^3(n)}{t_i^2(n)}\,(1+o(1)) & i=2,3,\ldots,K_n \\ -\,(L_n-1)\delta_5(n)(1+o(1)) & i=K_n+1 \end{cases}$$

and $\delta_5(n) = -(L_n/np_n)^{5/2}(2\log(1/p_n))^{1/2}$. Thus, we have

$$S_1(n) = \frac{[2(L_n-1)\tilde{T}'(n)\tilde{T}(n) + e'(n)\tilde{T}(n)]^2}{[4(L_n-1)^2\tilde{T}'(n)\tilde{T}(n) + 4(L_n-1)\tilde{T}'(n)e(n) + e'(n)e(n)\tilde{T}'(n)\tilde{T}(n)}$$

and, after some simplification,

$$1 - S_1(n) = \frac{[e'(n)e(n)\tilde{T}'(n)\tilde{T}(n) - [e'(n)\tilde{T}(n)]^2}{4n^2}\quad(1+o(1))$$

since $4(L_n-1)^2[\tilde{T}'(n)\tilde{T}(n)]^2 = 4n^2(1+o(1))$ asymptotically dominates the denominator of $S_1(n)$. Now,

$$\begin{aligned} e'(n)e(n) &= 2(L_n-1)^2\delta_5^2(n) + \frac{L_n^4(L_n-1)^2}{9n^4}\sum_{i=2}^{K_n}\frac{T_i^6(n)}{t_i^4(n)} \\ &= \frac{4L_n^7\log(1/p_n)}{n^5p_n^5} + \frac{1}{9}\,\frac{L_n^5}{n^5}\int_{p_n}^{1-p_n}\frac{T^6(x)}{t^4(x)}\,dx\ (1+o(1)). \end{aligned}$$

Using the method used in deriving (C.5) to evaluate the above integral,

we obtain

$$\int_{p_n}^{1-p_n} \frac{T^6(x)}{t^4(x)}\,dx = \frac{4}{3}\,\frac{\log(1/p_n)}{p_n^3} - \frac{5}{9p_n^3} + o(p_n^{-3}),$$

so that

$$e'(n)e(n) = \frac{4L_n^5 \log(1/p_n)}{27n^3 p_n^3}\,(1+o(1)). \tag{3.22}$$

Similarly,

$$e'(n)\tilde{T}(n) = \frac{4}{3}\,\frac{L_n^2}{np_n}\log(1/p_n)\,(1+o(1)). \tag{3.23}$$

Hence, since $\lim_{x\to 0} x\log x = 0$,

$$1 - S_1(n) = \frac{L_n^4}{27n^4 p_n^3}\log(1/p_n)\,(1+o(1)).$$

This completes the proof of lemma 3.5. QED

Lemma 3.6. The quantity $S_2(n)$ is asymptotically normally distributed with mean 0 and variance $(2L_n^4/(27n^5 p_n^3))(1+o(1))$.

Proof. Since $(\tilde{Y} - \tilde{T}(n))$ is asymptotically normal with mean vector 0, we need only determine the variance to complete the proof. Using $y = (y_1,\ldots,y_{K_n+1})$ as a dummy variable and letting $D(y)$ denote

$$\sum_{i=1}^{K_n+1} (y_i - y_\cdot)^2 ,$$

we have

$$\tilde{w}_i^{(1)}(\tilde{y}) = \frac{1}{\gamma'(n)\gamma(n)} \frac{\partial}{\partial y_i} \left[\frac{y'\gamma(n)\gamma'(n)y}{D(y)}\right]\Big|_{\tilde{y}}.$$

Using the facts that $\frac{\partial}{\partial y_i}(y'\gamma(n)\gamma'(n)y) = 2\gamma_i(n)\gamma'(n)y$ and $\frac{\partial}{\partial y_i}D(y) = 2(y_i-\bar{y})$, the above expression becomes

$$\tilde{w}_i^{(1)}(\tilde{y}) = \frac{2}{\gamma'(n)\gamma(n)}\left[\frac{\gamma_i(n)\gamma'(n)\tilde{y}}{D(\tilde{y})} - \frac{(\tilde{y}'\gamma(n)\gamma'(n)\tilde{y})(\tilde{y}_i - \tilde{y}.)}{D^2(\tilde{y})}\right] \tag{3.24}$$

In particular,

$$\tilde{w}_i^{(1)}(\tilde{T}(n)) = \frac{2}{\gamma'(n)\gamma(n)}\left[\frac{\gamma'(n)\tilde{T}(n)}{\tilde{T}'(n)\tilde{T}(n)}\right]\left[\gamma_i(n) - \frac{\gamma'(n)\tilde{T}(n)}{\tilde{T}'(n)\tilde{T}(n)}\tilde{T}_i(n)\right],$$

and using lemma 3.4, we can write the above expression for asymptotic purposes as

$$\tilde{w}_i^{(1)}(\tilde{T}(n)) = \frac{1}{n}\left(\gamma_i(n) - \frac{\gamma'(n)\tilde{T}(n)}{\tilde{T}'(n)\tilde{T}(n)}\tilde{T}_i(n)\right)(1+o(1)). \tag{3.25}$$

Since $\gamma'(n)\tilde{T}(n)/\tilde{T}'(n)\tilde{T}(n) \sim 2(L_n-1)$ and $\gamma_i(n) \sim 2(L_n-1)\tilde{T}_i(n)$ as $n \to \infty$, we must be careful about the smaller order terms in applying lemma 3.4 to the quantity in parentheses in (3.25). Recalling expressions (3.21) and (3.23), we can write

$$\tilde{w}_i^{(1)}(\tilde{T}(n)) = \begin{cases} \frac{L_n}{n}\delta_5(n) - \delta_6(n)\tilde{T}_1(n) & i=1 \\ \frac{1}{3}\frac{L_n^2(L_n-1)}{n^3}\frac{T_i^3(n)}{t_i^2(n)} - \delta_6(n)T_i(n) & i=2,3,\ldots,K_n \\ -\frac{L_n}{n}\delta_5(n) - \delta_6(n)\tilde{T}_{K_n+1}(n) & i=K_n+1 \end{cases} \tag{3.26}$$

where $\delta_6(n) = (4L_n^3/3n^3 p_n)\log(1/p_n)$. Noting that $\tilde{T}_1(n) = -q_n(2\log(1/p_n))^{1/2}$ as $n \to \infty$ and using (3.2), $(L_n/n)\delta_5(n) = o(\delta_6(n)\tilde{T}_1(n))$, so that the terms involving $\delta_5(n)$ in (3.26) can be neglected for asymptotic purposes, and the variance of $S_2(n)$ is given by the sum of the following twelve expressions:

$$4q_n^4(\log 1/p_n)\delta_6^2(n)\frac{L_n}{n(L_n-1)}\frac{p_n(1-p_n)}{t_1^2(n)} \tag{3.27}$$

$$-4q_n^4(\log 1/p_n)\delta_6(n)\frac{L_n}{n(L_n-1)}\frac{p_n^2}{t_1^2(n)} \tag{3.28}$$

$$\frac{2}{3}q_n^2(2\log 1/p_n)^{1/2}\delta_6(n)\frac{L_n^3 p_n}{n^4 t_1(n)}\sum_{j=2}^{K_n}\frac{T_j^3(n)}{t_j^3(n)}(1-p_n-(j-1)\frac{L_n}{n}) \tag{3.29}$$

$$-\frac{2}{3}q_n^2(2\log 1/p_n)^{1/2}\delta_6(n)\frac{L_n^3 p_n}{n^4 t_1(n)}\sum_{i=2}^{K_n}\frac{T_i^3(n)}{t_i^3(n)}(p_n + (i-1)\frac{L_n}{n}) \tag{3.30}$$

$$\frac{2}{3}q_n^2(2\log 1/p_n)^{1/2}\delta_6^2(n)\frac{L_n p_n}{n(L_n-1)t_1(n)}\sum_{j=2}^{K_n}\frac{T_j(n)}{t_j(n)}(1-p_n-(j-1)\frac{L_n}{n}) \tag{3.31}$$

$$-\frac{2}{3}q_n^2(2\log 1/p_n)^{1/2}\delta_6^2(n)\frac{L_n p_n}{n(L_n-1)t_1(n)}\sum_{i=2}^{K_n}\frac{T_i(n)}{t_i(n)}(p_n+(i-1)\frac{L_n}{n}) \tag{3.32}$$

$$\frac{1}{9}\frac{L_n^5(L_n-1)}{n^7}\sum_{i=2}^{K_n}\frac{T_i^6(n)}{t_i^6(n)}(p_n+(i-1)\frac{L_n}{n})(1-p_n-(i-1)\frac{L_n}{n}) \tag{3.33}$$

$$-\frac{2}{3}\delta_6(n)\frac{L_n^3}{n^4}\sum_{i=2}^{K_n}\frac{T_i^4(n)}{t_i^4(n)}(p_n+(i-1)\frac{L_n}{n})(1-p_n-(i-1)\frac{L_n}{n}) \tag{3.34}$$

$$\delta_6^2(n)\frac{L_n}{n(L_n-1)}\sum_{i=2}^{K_n}\frac{T_i^2(n)}{t_i^2(n)}(p_n+(i-1)\frac{L_n}{n})(1-p_n-(i-1)\frac{L_n}{n}) \tag{3.35}$$

$$\frac{2}{9}\frac{L_n^5(L_n-1)}{n^7}\sum_{i=2}^{K_n-1}\sum_{j=i+1}^{K_n}\frac{T_i^3(n)T_j^3(n)}{t_i^3(n)t_j^3(n)}(p_n+(i-1)\frac{L_n}{n})(1-p_n-(j-1)\frac{L_n}{n}) \tag{3.36}$$

$$\frac{-L_n^3}{3n^4}\delta_6(n)\sum_{i=2}^{K_n-1}\sum_{j=i+1}^{K_n}\left[\frac{T_i(n)T_j^3(n)}{t_i(n)t_j^3(n)}+\frac{T_i^3(n)T_j(n)}{t_i^3(n)t_j(n)}\right](p_n+(i-1)\frac{L_n}{n})(1-p_n-(j-1)\frac{L_n}{n}) \tag{3.37}$$

$$\frac{L_n}{n(L_n-1)}\delta_6^2(n)\sum_{i=2}^{K_n-1}\sum_{j=i+1}^{K_n}\frac{T_i(n)T_j(n)}{t_i(n)t_j(n)}(p_n+(i-1)\frac{L_n}{n})(1-p_n-(j-1)\frac{L_n}{n}). \tag{3.38}$$

We will now show that expression (3.36) asymptotically dominates the other terms in the above variance. For large n, expression (3.36) is approximately

$$\frac{2}{9}\frac{L_n^3(L_n-1)}{n^5}\int_{p_n}^{1-p_n}\int_{x}^{1-p_n} x(1-y)\frac{T^3(x)T^3(y)}{t^3(x)t^3(y)}dydx.$$

Evaluating the inner integral using (C.5) this becomes

$$\frac{1}{9}\frac{L_n^3(L_n-1)}{n^5} - \int_{p_n}^{1-p_n} x(1-x)\frac{T^5(x)}{t^5(x)}dx + \frac{2}{p_n}\int_{p_n}^{1-p_n} x\frac{T^3(x)}{t^3(x)}dx - \int_{p_n}^{1-p_n} x\frac{T^4(x)}{t^4(x)}dx + 0(p_n^{-2}),$$

which, again applying (C.5), is easily shown to equal

$$\frac{2}{27}\frac{L_n^4}{n^5p_n^3}(1+0(p_n)):$$

the first integral is zero by a symmetry argument; the second integral contributes $p_n^{-1}(T_1(n)/t_1(n))^2$ plus smaller order terms, and the third integral contributes $(1/3)(T_1(n)/t_1(n))^3$ plus smaller order terms.

Applying (C.7), we obtain the stated result.

Next we show that the remaining eleven expressions in the variance of $S_1(n)$ are $o(L_n^4/n^5 p_n^3)$ as $n \to \infty$.

Applying the definitions, lemma 3.2, and (C.10), we see that

$$[\text{expression } (3.27)] = \frac{32}{9} \frac{L_n^4}{n^5} \frac{(\log 1/p_n)^2}{p_n} (1+o(1))$$

$$= o(L_n^4/n^5 p_n^3)$$

as $n \to \infty$.

Expression (3.28) equeals p_n (expression (3.27)); so it is obviously of smaller order.

Adding expressions (3.29) and (3.30), applying definitions, approximating summation by integration and neglecting smaller order terms, we obtain the quantity

$$\frac{8}{9} \frac{L_n^4}{n^6} \int_{p_n}^{1-p_n} (1-2x) \frac{T^3(x)}{t^3(x)} dx.$$

Using (C.5) and (C.7) to evaluate the above integral, the expression becomes

$$- \frac{8}{9} \frac{L_n^4}{n^6 p_n^2} (1+o(1))$$

which is obviously of smaller order than (3.36) as $n \to \infty$.

Proceeding in exactly the same manner, we add (3.31) and (3.32) and obtain an expression of order $L_n^4(\log 1/p_n)^2/n^5$, which is again of smaller order than (3.36)

Approximating summation by integration and neglecting smaller order

terms, expression (3.33) becomes

$$\frac{1}{9}\,\frac{L_n^5}{n^6}\int_{p_n}^{1-p_n} x(1-x)\frac{T^6(x)}{t^6(x)}dx.$$

Integrating by parts twice and applying (C.5), the above expression is easily shown to equal

$$\frac{1}{18}\,\frac{L_n^5}{n^6 p_n^4}(1+O(p_n)),$$

and this is again of smaller order than (3.36).

Since (expression (3.34)) = $O(p_n \log(1/p_n)$expression (3.33)), it follows that (3.34) is of smaller order than (3.36).

Asymptotically, expression (3.35) is approximately

$$\frac{16}{9}\,\frac{L_n^6}{n^7 p_n^2}(\log 1/p_n)^2\int_{p_n}^{1-p_n} x(1-x)\frac{T^2(x)}{t^2(x)}dx,$$

which is $O(L_n^6(\log 1/p_n)^2/n^7 p_n^2)$ as $n \to \infty$. Thus, expression (3.35) is o(expression (3.36)).

Finally, since $(L_n^3/n^4 p_n^2)\delta_6(n)$ is $o(n^{-1}\delta_6^2(n))$ as $n \to \infty$, showing that expression (3.38) is of smaller order than (3.36) will imply the same for (3.37). But, (3.38) is approximately

$$\frac{16}{9}\,\frac{L_n^4}{n^5 p_n^2}(\log 1/p_n)^2\int_{p_n}^{1-p_n}\int_{x}^{1-p_n} x(1-y)\frac{T(x)T(y)}{t(x)t(y)}dydx.$$

Evaluating the inner integral this becomes

$$\frac{8}{9}\,\frac{L_n^4}{n^5 p_n^2}(\log 1/p_n)^2 \left\{ \begin{array}{l} [p_n T_1^2(n)+1-p_n-T_{K_n+1}(n)t_{K_n+1}(n)] \displaystyle\int_{p_n}^{1-p_n} x\frac{T(x)}{t(x)}dx \\[2ex] -\displaystyle\int_{p_n}^{1-p_n} x(1-x)\frac{T^3(x)}{t(x)}dx + \int_{p_n}^{1-p_n} xT^2(x)dx - \int_{p_n}^{1-p_n} x^2\,\frac{T(x)}{t(x)}dx \end{array} \right\}.$$

It is easily shown that the expression in braces is $O(1)$. Hence, (expression (3.38)) $= O(\frac{L_n^4}{n^5 p_n^3}p_n(\log 1/p_n)^2) = o($expression (3.36)) as $n \to \infty$. Summarizing, we have shown that the variance of $S_2(n)$ is given by $(2L_n^4/(27n^5p_n^3))(1+o(1))$, which completes the proof of lemma 3.6. QED

Lemma 3.7. The quantity $S_3(n) + S_4(n)$ is $o_p(L_n^2/n^{5/2}p_n^{3/2})$ as $n \to \infty$.

Proof. We first consider $S_3(n)$. Differentiating (3.24) with respect to y_j and simplifying terms, we obtain the following expression for the second partial derivatives of $\tilde{W}(n)$:

$$\tilde{w}_{ij}^{(2)}(\tilde{y}) = \frac{4}{\gamma'(n)\gamma(n)} \left\{ \begin{array}{l} \dfrac{\gamma_i(n)\gamma_j(n)}{2D(\tilde{y})} + \dfrac{2\tilde{y}'\gamma(n)\gamma'(n)\tilde{y}[\tilde{y}_i-\tilde{y}.][\tilde{y}_j-\tilde{y}.]}{D^3(\tilde{y})} \\[3ex] -\gamma'(n)y\dfrac{\gamma_i(n)[\tilde{y}_j-\tilde{y}.]+\gamma_j(n)[\tilde{y}_i-\tilde{y}.]+\frac{\tilde{y}'\gamma(n)}{2}[\delta_{ij}-(K_n+1)^{-1}]}{\tilde{D}^2(\tilde{y})} \end{array} \right\} \tag{3.39}$$

In particular,

$$\tilde{w}^{(2)}_{ij}(\tilde{T}(n)) = \frac{4}{\gamma'(n)\gamma(n)}\left\{\frac{\gamma_i(n)\gamma_j(n)}{2\tilde{T}'(n)\tilde{T}(n)} + \frac{2\tilde{T}_i(n)\tilde{T}_j(n)[\tilde{T}'(n)\gamma(n)]^2}{[\tilde{T}'(n)\tilde{T}(n)]^3} - \gamma'(n)\tilde{T}(n)\frac{\gamma_i(n)\tilde{T}_j(n)+\gamma_j(n)\tilde{T}_i(n)+\frac{\gamma'(n)\tilde{T}(n)}{2}[\delta_{ij}-(K_n+1)^{-1}]}{[\tilde{T}'(n)\tilde{T}(n)]^2}\right\}$$

Using the result of lemma 3.4, we can write the above expression for asymptotic purposes as

$$\tilde{w}^{(2)}_{ij}(\tilde{T}(n)) = 2(L_n/n)[(L_n/n)\tilde{T}_i(n)\tilde{T}_j(n) + (K_n+1)^{-1} - \delta_{ij}]. \quad (3.40)$$

Since

$$E\{S_3(n)\} = (1/2)\sum_{i=1}^{K_n+1}\sum_{j=1}^{K_n+1}\tilde{w}^{(2)}_{ij}(\tilde{T}(n))v_{ij}(n),$$

it is straightforward to show using (3.6), (3.40) and the methods of the proof of lemma 3.6 that the asymptotically dominating term in the above expectation is

$$-\frac{L_n^2}{n^2(L_n-1)}\sum_{i=2}^{K_n}\frac{(p_n+(i-1)\frac{L_n}{n})(1-p_n-(i-1)\frac{L_n}{n})}{t_i^2(n)}$$

which becomes

$$n^{-1}\int_{p_n}^{1-p_n} x(1-x)t^{-2}(x)dx\ (1+o(1)).$$

Evaluating the above integral, we find that

$$E\{S_3(n)\} = 2n^{-1}\log(1/p_n)\ (1+o(1)) \quad (3.41)$$

as $n \to \infty$.

Next, we obtain an upper bound on the standard deviation of $S_3(n)$. To do so, we use the fact that for any random variables $x_1, x_2, \ldots, x_n$,

$$\sigma\left(\sum_{i=1}^{n} X_i \right) \leq \sum_{i=1}^{n} \sigma(X_i). \tag{3.42}$$

Note that if X is normally distributed with mean 0 and variance σ^2,

$$E\{X^{2m}\} = (2m-1)(2m-3)\ldots(1)\sigma^{2m}$$

so that

$$\begin{aligned}
\mathrm{Var}\{(\tilde{Y}_i-\tilde{T}_i(n))(\tilde{Y}_j-\tilde{T}_j(n))\} &\leq E\{\tilde{Y}_i-\tilde{T}_i(n))^2(\tilde{Y}_j-\tilde{T}_j(n))^2\} \\
&\leq [E\{\tilde{Y}_i-\tilde{T}_i(n))^4\}E\{(\tilde{Y}_j-\tilde{T}_j(n))^4\}]^{1/2} \\
&= 3\mathrm{var}\{\tilde{Y}_i\}\mathrm{Var}\{\tilde{Y}_j\} \\
&\leq \mathrm{const}\begin{cases}
[L_n \log(1/p_n)]^{-2} & i,j \varepsilon \{1, K_n+1\} \\
[nL_n \log(1/p_n) t_j^2(n)]^{-1} & i \varepsilon \{1, K_n+1\} \\
 & j \not\varepsilon \{1, K_n+1\} \\
[nL_n \log(1/p_n) t_i^2(n)]^{-1} & i \not\varepsilon \{1, K_n+1\} \\
 & j \varepsilon \{1, K_n+1\} \\
[nt_i(n) t_j(n)]^{-2} & i,j \not\varepsilon \{1, K_n+1\}
\end{cases}
\end{aligned} \tag{3.43}$$

where we have used lemma 3.3 and the definition (3.6). Thus, from (3.40), (3.42), and (3.43)

$$\sigma\{S_3(n)\} \leqq \sum_{i=1}^{K_n+1} \sum_{i=1}^{K_n+1} |\tilde{w}_{ij}^{(2)}(\tilde{T}(n))| \sigma\{(\tilde{Y}_i - \tilde{T}_i(n))(\tilde{Y}_j - \tilde{T}_j(n))\}$$

$$= c_1[n\log(1/p_n)]^{-1} + c_2 L_n^{1/2} n^{-3/2} \log(1/p_n) + c_3 n^{-1}(\log 1/p_n)^2$$

$$+ c_4[np_n \log(1/p_n)]^{-1}$$

where c_1, c_2, c_3, and c_4 are $0(1)$ as $n \to \infty$. The asymptotically dominating term in the above expression is easily seen to be the term associated with c_4; that is, $\sigma\{S_3(n)\} \leqq c_4[np_n\log(1/p_n)]^{-1}(1+o(1))$ as $n \to \infty$. But, the latter expression can be written as

$$c_4 \frac{L_n^2}{n^{5/2} p_n^{3/2}} \frac{n^{3/2} p_n^{1/2}}{L_n^2} (\log 1/p_n)^{-1}(1+o(1)).$$

Applying (3.4), we see that $\sigma\{S_3(n)\} = o(L_n^2/n^{5/2}p_n^{3/2})$. Similarly, from (3.41), $E\{S_3(n)\}$ is also $o(L_n^2/n^{5/2}p_n^{3/2})$. Thus, the term $S_3(n)$ is $o_p(\tilde{\sigma}(n))$ as desired.

We now deal with the higher order terms in the expansion (3.19). Applying (3.42), the standard deviation of the kth order term (k=3,4,...) is less than or equal to

$$\sum_{i_1=1}^{K_n+1} \sum_{i_2=1}^{K_n+1} \cdots \sum_{i_k=1}^{K_n+1} |\tilde{w}_{i_1 i_2 \cdots i_k}^{(k)}(\tilde{T}(n))| \sigma\{\tilde{Y}_{i_1} - \tilde{T}_{i_1}(n)) \ldots (\tilde{Y}_{i_k} - \tilde{T}_{i_k}(n))\}.$$

We will find upper bounds on the quantities in the above summation which are sufficient to show that the sum is $o_p(L_n^2/n^{5/2}p_n^{3/2})$ as $n \to \infty$.

To simplify notation for the remainder of the proof we define

$$A(y) = \frac{[\gamma'(n)y]^2}{\gamma'(n)\gamma(n)} \qquad B(y) = \sum_{i=1}^{K_n+1} (y_i - y_\cdot)^2$$

$$A_i'(\tilde{y}) = \frac{\partial A(y)}{\partial y_i}\Big|_{\tilde{y}} \qquad B_i'(\tilde{y}) = \frac{\partial B(y)}{\partial y_i}\Big|_{\tilde{y}}$$

$$A_{ij}''(\tilde{y}) = \frac{\partial^2 A(y)}{\partial y_i \partial y_j}\Big|_{\tilde{y}} \qquad B_{ij}''(\tilde{y}) = \frac{\partial^2 B(y)}{\partial y_i \partial y_j}\Big|_{\tilde{y}}$$

for $i,j = 1,2,\ldots,K_n+1$. We note that all higher order derivatives of $A(y)$ and $B(y)$ are zero. We have shown previously using lemma 3.4 that

$$\left|\frac{A_i'(\tilde{T}(n))}{A(\tilde{T}(n))}\right| \leq \text{const}\ \frac{L_n}{n} (\log 1/p_n)^{1/2}$$

$$\left|\frac{B_i'(\tilde{T}(n))}{B(\tilde{T}(n))}\right| \leq \text{const}\ \frac{L_n}{n} (\log 1/p_n)^{1/2}$$

$$\left|\frac{A_{ij}''(\tilde{T}(n))}{A(\tilde{T}(n))}\right| \leq \text{const}\ \frac{L_n}{n} (\log 1/p_n)^{2} \tag{3.44}$$

$$\left|\frac{B_{ij}''(T(n))}{B(T(n))}\right| \leq \begin{cases} \text{const}(L_n/n) & \text{if } i=j \\ \text{const}(L_n/n)^2 & \text{if } i\neq j \end{cases}$$

as $n \to \infty$.

Now let $R(y) = \log(A(y)/B(y))$. It is easily seen that the general kth partial derivative of $R(y)$, which we will denote by $R^{(k)}_{i_1 i_2 \ldots i_k}(y)$,

involves only sums of products of the forms

$$\text{const}\left[\frac{A''_{j_1j_2}(y)}{A(y)}\right]\left[\frac{A''_{j_3j_4}(y)}{A(y)}\right]\cdots\left[\frac{A''_{j_{m_1-1}j_{m_1}}(y)}{A(y)}\right]\left[\frac{A'_{j_{m_1+1}}(y)}{A(y)}\right]\left[\frac{A'_{j_{m_1+m_2}}(y)}{A(y)}\right]$$

and

$$\text{const}\left[\frac{B''_{j_1j_2}(y)}{B(y)}\right]\left[\frac{B''_{j_3j_4}(y)}{B(y)}\right]\cdots\left[\frac{B''_{j_{m_3-1}j_{m_3}}(y)}{B(y)}\right]\left[\frac{B'_{j_{m_3+1}}(y)}{B(y)}\right]\left[\frac{B'_{j_{m_3+m_4}}(y)}{B(y)}\right]$$

where $m_1 + m_2 = m_3 + m_4 = k$. But, by applying the definitions and rearranging terms, we also have

$$\begin{aligned}
\tilde{w}_i^{(1)}(\tilde{T}(n)) &= S_1(n)R_i^{(1)}(\tilde{T}(n)) \\
\tilde{w}_{ij}^{(1)}(\tilde{T}(n)) &= S_1(n)\{R_i^{(1)}(\tilde{T}(n))R_j^{(1)}(\tilde{T}(n)) + R_{ij}^{(2)}(\tilde{T}(n))\} \\
\tilde{w}_{ijk}^{(3)}(\tilde{T}(n)) &= S_1(n)\left\{\begin{array}{l} R_i^{(1)}(\tilde{T}(n))R_{jk}^{(2)}(\tilde{T}(n)) + R_j^{(1)}(\tilde{T}(n))R_{ik}^{(2)}(\tilde{T}(n)) + \\ R_k^{(1)}(\tilde{T}(n))R_{ij}^{(2)}(\tilde{T}(n)) + R_{ijk}^{(3)}(\tilde{T}(n)) + \\ R_i^{(1)}(\tilde{T}(n))R_j^{(1)}(\tilde{T}(n))R_k^{(1)}(\tilde{T}(n)) \end{array}\right\} \\
&\vdots \\
\tilde{w}_{i_1i_2\ldots i_k}^{(k)}(\tilde{T}(n)) &= S_1(n)\left\{\begin{array}{l} \text{terms involving products of derivaties of} \\ R(y)\text{, where the sum of the orders of the} \\ \text{derivatives in each product is } k \text{ and each} \\ \text{index (i.e. } i_1,i_2,\ldots) \text{ is represented exactly} \\ \text{once in each product} \end{array}\right\}
\end{aligned} \tag{3.45}$$

Thus, since $S_1(n) \to 1$ as $n \to \infty$,

$$|\tilde{w}^{(k)}_{i_1 i_2 \ldots i_k}(\hat{T}(n))| \leq [\frac{L_n}{n}\log(1/p_n)]^{s^*} \tag{3.46}$$

where $s^* \equiv \max\{(k-s),[\frac{k+1}{2}]\}$

and s $\equiv$ the no. of <u>equalities</u> among the indices $i_1,i_2,\ldots,i_k$.

Next we turn to the standard deviations of the $(\tilde{Y}_{i_1},\ldots,\tilde{Y}_{i_k})$ terms in the expansion (3.19). The argument used in obtaining (3.43) also yields the following fact for $k=3,4,5\ldots$:

$$\mathrm{Var}\{(\tilde{Y}_{i_1}-\tilde{T}_{i_1}(n))(\tilde{Y}_{i_2}-\tilde{T}_{i_2}(n))\ldots(\tilde{Y}_{i_k}-\tilde{T}_{i_k}(n))\} \leq$$

$$\mathrm{Var}\{(\tilde{Y}_{i_1}-\tilde{T}_{i_1}(n))\ldots(\tilde{Y}_{i_{k-1}}-\tilde{T}_{i_{k-1}}(n))\cdot\text{const.}\begin{Bmatrix}[L_n\log(1/p_n)]^{-1} & i_k\varepsilon\{1,K_n+1\} \\ [np_n\log(1/p_n)]^{-1} & i_k\notin\{1,K_n+1\}\end{Bmatrix} \tag{3.47}$$

For given k, let $I_k(r,m,s,s^*)$ denote the set of integers r,m,s,s^* which satisfy

$$0 \leq m \leq k$$

$$\max\{0,m-2\} \leq r \leq \max\{m-1,0\}$$

$$r \leq s \leq r + \max\{k-m-1,0\}$$

$$s^* \equiv \max\{(k-s),[\frac{k+1}{2}]\}$$

Then, combining (3.46) and (3.47) we obtain

$$\sum_{i_1=1}^{K_n+1}\sum_{i_2=1}^{K_n+1}\cdots\sum_{i_k=1}^{K_n+1}|\tilde{w}^{(k)}_{i_1 i_2\ldots i_k}(\hat{T}(n))|\sigma\{(\tilde{Y}_{i_1}-\tilde{T}_{i_1}(n))\ldots(\tilde{Y}_{i_k}-\tilde{T}_{i_k}(n))\}$$

$$\leq \text{const} \cdot \max_{I_k(r,m,s,s^*)} \{(K_n+1)^{k-s-m+r} \left(\frac{L_n}{n}\right)^{s^*-(m/2)} (\log 1/p_n)^{s^*-(k/2)}\, n^{-k/2}\, p_n^{(m-k)/2}\}$$

for $k=3,4,\ldots$ (where m represents the number of indices $(i_1,\ldots,i_k)$ which assume values 1 or K_n+1; r represents the number of equalities among those indices, and s and s^* are as defined above). By separately considering each of the four cases: (1) k even, $s \geq k/2$; (2) k even, $s < k/2$; (3) k odd, $s \geq (k-1)/2$; and (4) k odd, $s < (k-1)/2$, it is straightforward to show that the right hand side above is asymptotically maximized by the following parameter settings:

$$\begin{cases} m=r=0,\ s=1,2,\text{or } 3 & \text{if } k=3 \\ m=r+1=s+1=\frac{k+1}{2} & \text{if } k>3 \text{ and odd} \\ m=r+1=s+1=\frac{k+2}{2} & \text{if } k>3 \text{ and even} \end{cases}$$

in which cases it becomes

$$\text{const.} \begin{cases} \dfrac{\log(1/p_n)^{1/2}}{n^3 p_n^3} & \text{if } k=3 \\[2ex] \left[\dfrac{\log(1/p_n)}{L_n^{k-3} n^{k+3} p_n^{k+1}}\right]^{1/4}, & \text{if } k>3 \text{ and odd} \\[2ex] [L_n^{k-2} p_n^{k-2} n^{k+2}]^{-1/4} & \text{if } k>3 \text{ and even} \end{cases} .$$

Successive terms of the above expression are decreasing in k at a rate which is $o(L_n^{-1/2})$, while for $k = 3$ we have

$$\left[\frac{\log(1/p_n)}{n^3 p_n^3}\right]^{1/2} = \frac{L_n^2}{n^{5/2} p_n^{3/2}} \frac{n(\log 1/p_n)^{1/2}}{L_n} = o\left(\frac{L_n^2}{n^{5/2} p_n^{3/2}}\right)$$

as n tends to infinity. Hence, a further application of (3.42) establishes that the standard deviation of $S_4(n)$ is $o(\tilde{\sigma}(n))$.

Using (3.46) and the formula analogous to (3.43) for expectations, it is easily shown in the same manner that the expected value of $S_4(n)$ is also $o(\tilde{\sigma}(n))$ as $n \to \infty$. Thus, $S_4(n)$ is $o_p(\tilde{\sigma}(n))$ as claimed. This completes the proof of lemma 3.7. QED

Combining the results of lemmas 3.5, 3.6, and 3.7, the proof of Theorem 3.1 is complete.

3.2.4 Discussion

In this subsection, we explain the reason for the definition (3.5) of the quantity q_n, discuss the possible relationship between $\tilde{W}(n)$ and W, including the introduction of a new statistic $\tilde{W}^*(n)$ which is an analog of W^* of Shapiro and Francia [36], and finally, mention ' ne practical considerations in choosing the parameters of our selected subset of order statistics.

The quantity q_n. In order to ensure the asymptotic normality of our selected subset of order statistics, it was necessary to have the "tail spacing" (np_n) of the statistics different from the "internal spacing" (L_n). A consequence of this fact was that the elements of the inverse of the asymptotic covariance matrix corresponding to the tail statistics were too large, upsetting the "balance" inherent in the coefficients $\gamma(n) = V^{-1}(n)T(n)$. Specifically, these coefficients were asymptotically dominated by $\gamma_1(n)$, and, in fact, $\gamma'(n)\gamma(n)$ was asymptotically dominated

by $Y_1^2(n)$. Thus, for $q_n \equiv 1$, not only would the value of the statistic $\tilde{W}(n)$ be essentially determined by the value of the tail statistics $Y_1(n)$ and $Y_{K_n+1}(n)$, but the statistic itself would converge to <u>zero</u> instead of one as $n \to \infty$, in which case the $\tilde{W}$-test would no longer be consistent. Even defining $q_n = (np_n/L_n)^{1/2}$, its limiting form, had the effect of allowing the tail terms to dominate both the rate of convergence of $S_1(n)$ to one and the variance of $S_2(n)$. Only by defining q_n as we did were we able to cause the effect of the tail statistics to be of smaller order than the combined effect of the remaining K_n-1 statistics.

<u>The relationship between $\tilde{W}$ and W.</u> Probably the most intriguing and certainly the most difficult question raised by the preceding work is whether or not we can infer anything about the large sample behavior of W from the asymptotic distribution of $\tilde{W}$ under the null hypothesis (or under the alternatives of section 3.4). In essence, does the form of the statistic $\tilde{W}$ properly encode the information about the null hypothesis in order that the asymptotic sufficiency result of Weiss [40] may be extended to this case? If we were to take p_n and L_n to be their values if the full set of order statistics were being considered (i.e. $p_n = 1/n$, $L_n = 1$), then the variance of $\tilde{W}(n)$ would be of order n^{-2}, which is what has been conjectured for W by Stephens [33] and others. However, we of course cannot do that, and, in fact, condition (3.4) precludes any choice of the parameters which would make the variance that small.

On the other hand, let us define an analog of the Shapiro-Francia statistic W^* based on $\tilde{Y}(n)$ (with $q_n \equiv 1$) by

$$\tilde{W}^*(n) = \frac{[T'(n)\tilde{Y}]^2}{[T'(n)T(n)]\tilde{S}^2(n)} . \tag{3.48}$$

Also, let a_n^* denote the quantity

$$\frac{L_n}{n} \sum_{i=1}^{K_n+1} \sum_{j=1}^{K_n+1} [(L_n/n)T_i(n)T_j(n) + (K_n+1)^{-1} - \delta_{ij}]v_{ij}(n)$$

where $\{v_{ij}(n)\}$ are given by (3.6) (again with $q_n = 1$). In the proof of lemma 3.7, a_n^* was shown to be $O(\log 1/p_n)$ as $n \to \infty$. In appendix B we prove the following theorem:

Theorem 3.2. Under the null hypothesis,

$$n[1 - \tilde{W}^*(n)] - a_n^* \tag{3.49}$$

converges in law to $\Psi(y)$, where $\Psi(y)$ is the distribution of

$$\sum_{k=3}^{\infty} \frac{(X_k^2 - 1)}{k}$$

with $X_3, X_4, \ldots$ i.i.d. standard normal random variables.

The significance of Theorem 3.2 is that the analog $W^*(n)$ of the Shapiro-Francia statistic, or, more precisely, of the De Wet-Venter r_n^2, has the same asymptotic distribution (under H_o and up to a shift in location) as the corresponding statistic based on the full set of order statistics. Thus, the form of $\tilde{W}^*(n)$ does properly encode the information about the

null hypothesis. We can conclude, then, that if $\hat{W}(n)$ does not behave like W, it is because of differences between the rate at which the asymptotic eigenvector result (lemma 3.4) takes hold for the mean vector and covariance matrix of $\tilde{Y}$ as opposed to the rate for the corresponding result for the full set of order statistics. Progress toward rigorously defining the relationship between $\tilde{W}$ and W might proceed along that line of analysis.

Practical considerations. In practice, there are two ways of choosing the "spacing" parameter L_n:

(A) $L_n = n^{\delta}$ where $2/3 < \delta <$ ' (such that (3.1) - (3.4) hold)

(B) $L_n = n^{r(n)}$ where $r(n) \to c$ as $n \to \infty$, $2/3 < c < 1$ (such that (3.1) - (3.4) hold).

For moderately large n, method (A) leads to larger values of L_n and, hence, smaller values of K_n than method (B). The advantage ethod (A) is that the asymptotic theory takes over for smaller n. dis-advantages of method (A) are: (1) K_n+1, the number of selected order statistics, is very small (for example, for $L_n = n^{3/4}$ and $n = 1000$, $(K_n+1) = 5$); and (2) the statistic $\hat{W}(n)$ is too sensitive to small deviations in the tail values. Table 3.1 gives the asymptotic means and standard deviations for $\hat{W}(n)$ for small changes in the tail percentiles p_n, $1-p_n$. Small changes in the actual values of the tail order statistics should also have dramatic effects.

Method (B) is preferable in that smaller values of L_n arise; that is, the asymptotic variance of $\tilde{W}(n)$ is smaller and we select a larger subset of the order statistics. The problem with this method is that even

TABLE 3.1

ASYMPTOTIC MEANS AND VARIANCES OF $\tilde{W}(n)$

FOR VARIOUS CHOICES OF THE TAIL QUANTILES

	CASE*					
	I		II		III	
n	$\tilde{\mu}(n)$	$\tilde{\sigma}(n)$	$\tilde{\mu}(n)$	$\tilde{\sigma}(n)$	$\tilde{\mu}(n)$	$\tilde{\sigma}(n)$
100	.503	.068	.909	.033	.956	.024
150	.544	.052	.928	.024	.967	.017
200	.575	.043	.940	.019	.973	.013
250	.599	.037	.947	.015	.977	.011
300	.619	.033	.953	.013	.980	.009
350	.635	.029	.957	.012	.982	.008
400	.650	.027	.961	.010	.984	.007
450	.662	.025	.964	.009	.986	.006
500	.673	.023	.966	.009	.987	.006
550	.683	.021	.968	.008	.988	.005
600	.692	.020	.970	.007	.988	.005
650	.700	.019	.972	.007	.989	.005
700	.707	.018	.973	.006	.990	.004
750	.714	.017	.974	.006	.990	.004
800	.720	.016	.975	.006	.991	.004
850	.726	.016	.976	.005	.991	.004
900	.732	0.15	.977	.005	.992	.003
950	.737	.015	.978	.005	.992	.003
1000	.741	.014	.979	.005	.992	.003

*Case	L_n	p_n
I	$n^{.75}$	$n^{-(1/\log n)^{1/2}}$
II	$n^{.75}$	$n^{-(1/\log n)^{2/3}}$
III	$n^{.75}$	$n^{-(1/\log n)^{3/4}}$

for moderate n, significant contributions to the mean and variance of $\tilde{W}(n)$ come from the higher order terms in the expansion (3.19) as well as the lower order components of the constant and linear terms.

From the practical point of view, therefore, the asymptotic theory is not overly useful for moderately large n in actually specifying the critical region of the test. It is useful in providing possible insight into the behavior of W, in indicating how to choose L_n and p_n to give the best results (i.e. choose L_n "small" and p_n "large"--which makes sense since both the numerator and denominator of $\tilde{W}(n)$ have interpretations as estimators of σ^2), and especially, in indicating how $\tilde{W}(n)$ behaves under alternative distributions so that comparisons with other test procedures may be made. The latter results are discussed in the remainder of this chapter.

3.3 Consistency of the $\tilde{W}$-Test

In this section we show that as $n \to \infty$, the power of the sequence of tests based on $\tilde{W}(n)$ tends to one against any fixed alternative. Throughout this section we assume that for each n, the sample $X_1,\ldots,X_n$ comes from the distribution $H(x)$, where $H(x) = \Phi((x-\mu)/\sigma)$ for any (μ,σ) with $\sigma > 0$. In particular, we assume that there exists some $\varepsilon > 0$ such that

$$\inf_{\substack{c,d \\ c>0}} \sup_{0<p<1} |H^{-1}(p) - c\Phi^{-1}(p) - d| > \varepsilon \qquad (3.50)$$

where $H^{-1}(p)$ is defined as the smallest value of x such that $H(x) \geq p$. (Note that the degenerate case where the support of $H(x)$ is a single point

is ruled out by this assumption.) Defining

$$\eta(x) = \begin{cases} |x|^{\delta}H(x) & \text{if } x \leq 0 \\ x^{\delta}(1-H(x)) & \text{if } x > 0 \end{cases},$$

we further <u>assume</u> that $\eta(x)$ is bounded and converges to zero as $|x| \to \infty$ for some $\delta > 0.4$. We remark at the end of this section how this regularity condition may be relaxed. A sufficient condition that it hold (see, e.g. Sen [23]) is that $E_H\{|x|^{\delta}\}$ is finite. Our proof of consistency requires the following lemma:

<u>Lemma 3.8</u>. <u>There are points</u> p_1^*, p_2^*, p_3^*, <u>with</u> $0 < p_i^* < 1$ <u>for</u> $i=1,2,3$, <u>non-overlapping open intervals</u> $I_i^* = (p_i^*-\delta^*, p_i^*+\delta^*)$ <u>with</u> $\delta^* > 0$, <u>and some</u> $\varepsilon^* > 0$ <u>such that for any</u> (c,d) <u>with</u> $c > 0$,

$$\inf_{p \in I_i^*} |H^{-1}(p) - c\Phi^{-1}(p) - d| > \varepsilon^* \tag{3.51}$$

<u>for at least one</u> $i=1,2,3$.

<u>Proof</u>. We prove the result when $H(x)$ is discrete. The modifications required when $H(x)$ is continuous are pointed out parenthetically as we proceed.

Let x_1 and x_2 $(x_1 < x_2)$ be any two jump points of $H(x)$ and choose p_i^*, $i=1,2$, such that $H(x_i-) < p_i^* < H(x_i)$, where $H(x-)$ denotes the limit from the left. (In the continuous case, we choose p_i^* such that $0 < p_1^* < p_2^* < 1$ and $H^{-1}(p)$ is strictly increasing for p in some neighborhood of p_i^*.) Let $\hat{c}$ and $\hat{d}$ be the uniquely determined

values such that $\Phi((x_i-\hat{d})/\hat{c}) = p_i^*$ for $i=1,2$. Then, from (3.50), there exists p_3^*, $0 < p_3^* < 1$, such that

$$|H^{-1}(p_3^*) - \hat{c}\Phi^{-1}(p_3^*) - \hat{d}| > \varepsilon .$$

Furthermore, the continuity of $\hat{c}\Phi^{-1}(p) + \hat{d}$ implies that we can "shift" p_3^* slightly if necessary so that $H(x_3-) < p_3^* < H(x_3)$ for some jump point x_3 of $H(x)$, possibly at the expense of replacing ε by a smaller positive quantity ε' (say). (In the continuous case, we can have $H^{-1}(p)$ strictly increasing in a neighborhood of p_3^*.)

Now consider

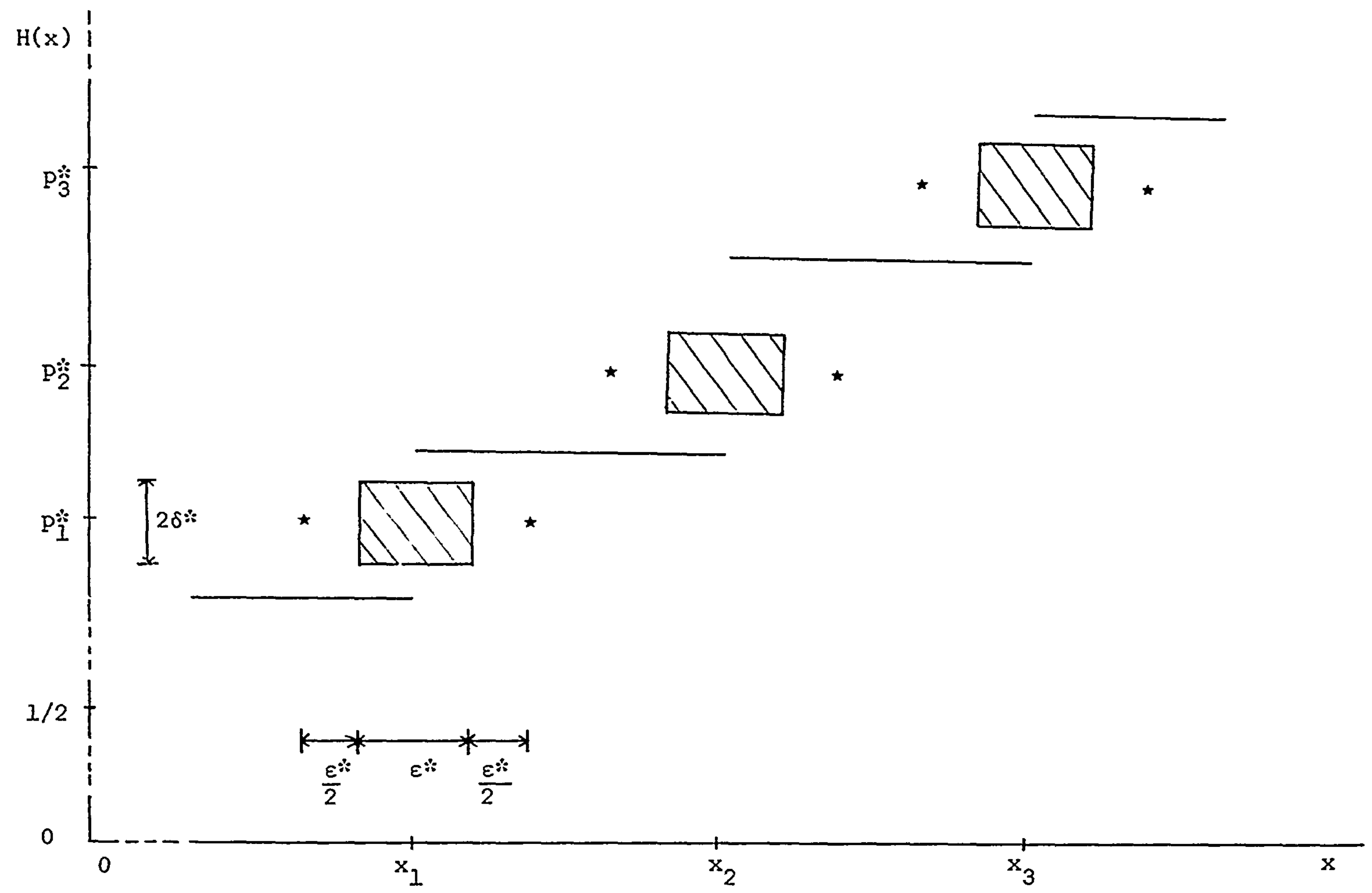
$$\inf_{\substack{c,d \\ c>0}} \max_{i=1,2,3} |H^{-1}(p_i^*) - c\Phi^{-1}(x_i) - d| .$$

By the way the p_i^* were chosen, the value of this quantity must be strictly positive. We will denote the value by $2\varepsilon^*$.

We now show how the intervals I_i^* are determined. We consider only the case where $x_1 < x_2 < x_3$ and $1/2 < p_1^* < p_2^* < 1$. The other cases are similar. The situation is depicted in Figure 3.2. The shaded boxes are the regions where (3.51) is satisfied; that is, (3.51) holds whenever the curve $\Phi((x-d)/c)$ passes through no more than two of the three boxes. By the definition of ε^*, any normal curve must pass through or outside at least one of the points $(x_i \pm \varepsilon^*, p_i^*)$, $i=1,2$, or 3. These points are denoted by $\star$ in the figure. Since the slope of $\Phi((x-d)/c)$ is decreasing in x, the worst possible case is that c and d satisfy

FIGURE 3.2

LEMMA 3.8 IN THE DISCRETE CASE



$$\Phi\left(\frac{x_1 - (\varepsilon^*/2) - d}{c}\right) = p_1^* + \delta^*$$

$$\Phi\left(\frac{x_3 - \varepsilon^* - d}{c}\right) \quad = p_3^*$$

Adding the additional equation

$$\Phi\left(\frac{x_3 - (\varepsilon^*/2) - d}{c}\right) = p_3^* + \frac{3\delta^*}{2}$$

solving for δ^*, and taking $\delta_i^* = \delta^*$, $i=1,2,3$, thus yields the desired intervals. (In the continuous case, the shaded boxes become cylinders of width ε^* about the curve $H(x)$ and the slope of $H(x)$ in each of the neighborhoods must, therefore, also enter the calculation, but the method is the same.) This completes the proof of lemma 3.8. QED

Now let $J_i^*(n)$, $i=1,2,3$, denote the set

$$\{j \mid p_i^* - \delta^*/2 \leqq p_n + (j-1)L_n/n \leqq p_i^* + \delta^*/2\}.$$

Note that $|J_i^*(n)|$ = largest integer $\leqq \delta^* n/L_n$. We now show that (3.51) implies that for at least one $i=1,2,3$,

$$\lim_{n\to\infty} P\{\min_{j\varepsilon J_i^*(n)} |\tilde{Y}_j - c\tilde{T}_j(n) - d| \geqq \varepsilon^*\} = 1 \tag{3.52}$$

First assume that $H(x)$ is discrete. By the way the p_i^* were chosen, we may assume without loss of generality that $H(x_i-) < p_i^* - \delta^*/2 < p_i^* + \delta^*/2 <$ $H(x_i)$ (since otherwise we just replace I_i^* by its intersection with $(H(x_i-),H(x_i)))$. Thus, $j\varepsilon\ J_i^*(n)$ implies $Y_j = x_i$ with probability

approaching one as $n \to \infty$. Thus, (3.52) follows immediately from (3.51) for this case. Now assume that $H(x)$ is continuous. We may again assume without loss of generality that $H^{-1}(p)$ is strictly increasing on I_i^*. Let $M_{ij}(n)$ denote the quantity

$$\sqrt{n}\left(H(\tilde{Y}_j) - \frac{np_n + (j-1)L_n}{n}\right) \qquad j \in J_i^*(n).$$

It is well known that $\max_{j \in J_i^*(n)} |M_{ij}(n)|$ is finite with probability one. Writing

$$\tilde{Y}_j = H^{-1}\left(\frac{M_{ij}(n)}{\sqrt{n}} + \frac{np_n+(j-1)L_n}{n}\right),$$

we see that $j \in J_i^*(n)$ implies that the argument of H^{-1} above is in I_i^* with probability approaching one as n increases. Applying (3.51), we again obtain (3.52).

For $i=1,2,\ldots,K_n+1$, let $\gamma_i^*(n)$ denote the quantity $\gamma_i(n)/[\gamma'(n)\gamma(n)]^{1/2}$ (where $\gamma(n) = \tilde{T}'(n)V^{-1}(n)$ as previously defined). Consider the quantity

$$\sum_{j=1}^{K_n+1} [\gamma_j^* - u_n(\tilde{Y}_j - \tilde{Y}.)]^2 \tag{3.53}$$

where u_n is defined as

$$\left(\sum_{j=1}^{K_n+1} \gamma_j^*\tilde{Y}_j\right)\Big/\left(\sum_{j=1}^{K_n+1} (\tilde{Y}_j - \tilde{Y}.)^2\right). \qquad (3.53)$$

It is easily seen that (3.53) is equal to $1 - \tilde{W}(n)$. In lemma 3.4 it was shown that $\gamma_j^*(n) = \tilde{T}_j(n)(1+0(L_n/np_n))(K_n+1)^{-1/2}$; hence (3.53) can also

be written as

$$u_n^2 \sum_{j=1}^{K_n+1} [\tilde{T}_j(n)(1+O(L_n/np_n))(u_n^2(K_n+1)^{-1/2}) - \tilde{Y}_i + \tilde{Y}.]^2.$$

Applying lemma 3.8,

$$1 - \tilde{W}(n) \geqq u_n^2 \sum_{i=1}^{3} \sum_{j\varepsilon J_i^*} [\tilde{T}_j(n)(1+O(L_n/np_n))(u_n^2(K_n+1)^{-1/2} - \tilde{Y}_j + \tilde{Y}.]^2$$

$$\geqq \delta^*(\varepsilon^*)^2(nu_n^2/L_n)$$

with probability approaching 1 as $n \to \infty$. Also,

$$(nu_n^2/L_n) > (K_n+1)u_n^2 = \tilde{W}(n)/\tilde{S}^2(n)$$

with $\tilde{S}^2(n)$ as defined in section 3.2. Thus,

$$1 - \tilde{W}(n) \geqq \frac{\delta^*(\varepsilon^*)^2}{\tilde{S}^2(n) + \delta^*(\varepsilon^*)^2} \qquad (3.54)$$

with probability approaching one as $n \to \infty$. Since the sequence of critical points for the $\tilde{W}$-test approaches one at the rate $n^4p_n^3/(L_n^4\log(1/p_n))$ (see lemma 3.5 and Figure 3.1), consistency will follow immediately once we establish that $\tilde{S}^2(n)$ is $o_p(n^4p_n^3/(L_n^4\log(1/p_n)))$ as $n \to \infty$.

We now return to the function $\eta(x)$. Our initial assumption implies that $\eta(\tilde{Y}_i)$ is bounded for $1 \leq i \leq K_n+1$. Using this fact together with the fact that $\sqrt{n}(H_n(\tilde{Y}_j) - H(\tilde{Y}_i))$ is finite with probability one (where $H_n(x)$ is the empirical c.d.f. and $H(x)$ may be discrete or continuous),

it follows that the function

$$\begin{cases} |\tilde{Y}_i|^{\delta}(p_n+(i-1)L_n/n) & \text{if } p_n+(i-1)L_n/n \leq H(0) \\ |\tilde{Y}_i|^{\delta}(1-p_n-(i-1)L_n/n) & \text{if } p_n+(i-1)L_n/n > H(0) \end{cases} \tag{3.55}$$

is finite with probability one. If $H(x)$ is continuous, (3.55) is immediately implied by the fact that $n^{-1/2}p_n^{-1}$ converges to zero as $n \to \infty$. If $H(x)$ is discrete and $\tilde{Y}_i \leq 0$, then $H_n(\tilde{Y}_i) \leq p_n+(i-1)L_n/n$, from which the first half of (3.55) follows. Recalling our assumption on the tails of $H(x)$, for $\tilde{Y}_i > 0$ and $H(x)$ discrete, $1-p_n-(i-1)L_n/n = H_n(\tilde{Y}_i) + \delta_7(n)$, where $\delta_7(n)$ is $O(\tilde{Y}_i)$ and, in fact, converges to zero as the sample quantile $\tilde{Y}_i$ assumes arbitrarily large values. Thus, (3.55) is established, and for <u>any</u> $\beta > 0$, we can find a constant $M(\beta)$ such that $P_H\{$(expression (3.55)) $< M(\beta)\} > 1 - \beta$.

Finally, with probability greater than $1 - \beta$,

$$\tilde{S}^2(n) \leq (K_n+1)^{-1} \sum_{i=1}^{K_n+1} \tilde{Y}_i^2$$

$$< \frac{M(\beta)}{K_n+1} \sum_{i=1}^{i^*} [p_n+(i-1)L_n/n]^{-2/\delta} + \sum_{i=i^*+1}^{K_n} [1-p_n-(i-1)L_n/n]^{-2/\delta}$$

(where i^* is defined by $p_n+(i^*-1)L_n/n < H(0) < p_n+i^*L_n/n$)

$$< 2\frac{M(\beta)}{K_n+1} \sum_{i=1}^{i^*+1} [p_n+(i-1)L_n/n]^{-2/\delta}$$

$$< 2M(\beta)p_n^{-2/\delta} .$$

But, using (3.2),

$$\frac{L_n^4 \log(1/p_n)}{n^4 p_n^3} p_n^{-2/\delta} = \frac{L_n^4}{n^4 p_n^8 (\log 1/p_n)^4} p_n^{5-(2/\delta)} (\log 1/p_n)^5$$

which converges to zero for $\delta > 0.4$. Thus, since β can be made arbitrarily small, $\tilde{S}^2(n)$ is $o_p(n^4 p_n^3/(L_n^4 \log(1/p_n)))$ as required.

Summarizing the above, we have proved

Theorem 3.3. Under the above assumptions, the $\tilde{W}$-test is consistent.

We remark that the same proof immediately yields the consistency of the test based on $\tilde{W}^*$. The only change above is that $\gamma_i^*(n) \equiv T_i(n)$. Also, the restriction $\delta > 0.4$ can be relaxed by placing tighter controls on the tail quantiles (p_n and $1-p_n$). For example, if instead of (3.2) we assume that $\lim_{n\to\infty} L_n/np_n^4 = 0$, then theorem 3.3 follows for $\delta > 2/9$, and so forth.

3.4 The Asymptotic Power of the $\tilde{W}$-Test

We wish to develop a means of comparing the $\tilde{W}$-test with other known tests for normality. As in the previous chapter, we do so by defining a general class of alternatives which approaches $\Phi(x)$ as n increases and determining the exact rate and character of approach which guarantees a non-trivial power.

The sequence of alternative distributions which we will study has the form

$$H_n(x) = \Phi(x) + \varepsilon_n(x) \tag{3.56}$$

where the disturbance functions $\varepsilon_n(x)$ satisfy the following conditions:

$$\varepsilon_n''(x) \text{ exists and } |\varepsilon_n''(x)| < D < \infty \quad \text{for all} \quad x \in (-\infty,\infty) \tag{3.57}$$

$$-\Phi(x) \leqq \varepsilon_n(x) \leqq 1 - \Phi(x) \quad \text{for all} \quad x \in (-\infty,\infty) \tag{3.58}$$

$$\int_{-\infty}^{\infty} \varepsilon_n'(x)dx = 0 \tag{3.59}$$

$$\varepsilon_n'(x) \geqq -\phi(x) \qquad \text{for all} \quad x \in (-\infty,\infty). \tag{3.60}$$

The above definition and assumptions are (2.32) - (2.37) specialized to the normal case.

Denote by $\rho(n,\varepsilon_n)$ the quantity

$$\int_{p_n}^{1-p_n} [\frac{\varepsilon_n(T(x))}{t(x)}]^2 dx - [\int_{p_n}^{1-p_n} \frac{\varepsilon_n(T(x))}{t(x)} dx]^2 - [\int_{p_n}^{1-p_n} \frac{T(x)}{t(x)}\varepsilon_n(T(x))dx]^2$$

$$+ \frac{1}{3}(\frac{L_n}{n})^2 \{ \int_{p_n}^{1-p_n} T^2(x)dx \int_{p_n}^{1-p_n} [\frac{T(x)}{t(x)}]^3 \varepsilon_n(T(x))dx - \int_{p_n}^{1-p_n} \frac{T^4(x)}{t^2(x)}dx \int_{p_n}^{1-p_n} \frac{T(x)}{t(x)}\varepsilon_n(T(x))dx\}.$$

Then, in order to guarantee a power of the $\tilde{W}$-test between the level α and 1, we require:

$$\lim_{n\to\infty} \frac{n^{5/2}p_n^{3/2}}{L_n^2} \rho(n,\varepsilon_n) = d^* \tag{3.61}$$

for some constant d^*, $0 < d^* < \infty$, and, letting

$$C(n) \equiv \max\{ \frac{L_n^2(\log 1/p_n)^3}{np_n}, \frac{n^{3/2}p_n^{3/2}\log(1/p_n)}{L_n} \},$$

$$\lim_{n\to\infty} b(n) \sup_{p_n<x<1-p_n} \left|\frac{\varepsilon_n(T(x))}{t(x)}\right| = 0 \tag{3.62}$$

$$\lim_{n\to\infty} b(n)(\log 1/p_n)^{-1} \sup_{p_n<x<1-p_n} \left|\frac{\varepsilon_n'(T(x))}{t(x)}\right| = 0 \tag{3.63}$$

for some increasing sequence $b(n)$ such that

$$\lim_{n\to\infty} b(n) = \infty, \quad \lim_{n\to\infty} C(n)b^{-2}(n) = 0.$$

The motivation for the condition (3.61) will become evident in the course of the derivation of the asymptotic power of the test. The quantity $\rho(n,\varepsilon_n)$ is what might be termed the "characteristic metric" for the test; the relation (3.61) requires that the sequence $\{H_n(x)\}$ converge to $\Phi(x)$ according to this metric at a rate proportional to the asymptotic standard deviation of $\tilde{W}(n)$ under the null hypothesis. Heuristically, the conditions (3.62) and (3.63) insure that the alternative distributions do not oscillate too badly in the interval $(p_n,1-p_n)$. When (3.61) holds, (3.62) and (3.63) will also hold in most cases, although an additional restriction on the parameters (K_n,L_n,p_n) may result.

The remainder of this section is devoted to proving the following theorem:

Theorem 3.4. Under the sequence of distributions (3.56), the asymptotic power of the $\tilde{W}$-test is given by $\Phi((27/2)^{1/2}d^* + c_{\Phi,\alpha})$.

We divide the proof into two lemmas. The first introduces a sequence of distributions $\{\bar{H}_n(x)\}$ which will simplify computations in the sequel. The second contains the asymptotic distribution of $W(n)$ under the sequence

of alternatives.

Denote the function $T(x) - \varepsilon_n(T(x))/t(x)$ $(p_n \leqq x \leqq 1-p_n)$ by $\overline{B}_n(x)$, with derivative

$$\overline{B}_n'(x) = t^{-1}(x)[1 - \frac{\varepsilon_n(T(x))T(x)}{t(x)} - \frac{\varepsilon_n'(T(x))}{T(x)}].$$

Also, let $\overline{A}_n(x)$ denote the function given by

$$\overline{A}_n(x) = \begin{cases} \overline{a}_{1n}(x) & 0 < x < p_n \\ \overline{B}_n(x) & p_n < x < 1-p_n \\ \overline{a}_{2n}(x) & 1-p_n < x < p_n \end{cases}$$

where $\overline{a}_{1n}(x)$ and $\overline{a}_{2n}(x)$ are continuous, differentiable functions which satisfy

$$\overline{a}_{1n}(p_n) = \overline{B}_n(p_n) \qquad \overline{a}_{2n}(1-p_n) = \overline{B}_n(1-p_n)$$

$$\overline{a}_{1n}'(p_n) = \overline{B}_n'(p_n) \qquad \overline{a}_{2n}'(1-p_n) = \overline{B}_n'(1-p_n)$$

$$\lim_{x\to 0} \overline{a}_{1n}(x) = -\infty \qquad \lim_{x\to 1} \overline{a}_{2n}(x) = +\infty$$

$\overline{a}_{1n}(x)$ $(\overline{a}_{2n}(x))$ is strictly increasing on $(0,p_n)$ ($(1-p_n,1)$). We will not need explicit expressions for $\overline{a}_{1n}(x)$ and $\overline{a}_{2n}(x)$, but it is clear that such functions exist.

Thus, $\overline{A}_n(x)$ approaches $-\infty$ as $x \to 0$ and $+\infty$ as $x \to 1$. Furthermore, by (3.62) and (3.63), there exists some N_o such that for all

$n > N_o$, $\overline{B}_n(x)$ (and hence $\overline{A}_n(x)$) is strictly increasing on $[p_n, 1-p_n]$ $((0,1))$. Since we are concerned only with asymptotic results, we lose no generality by assuming that $n > N_o$, and we do so from now on. In particular, we can define an absolutely continuous c.d.f. $\overline{H}_n(y)$ by specifying its inverse:

$$\overline{H}_n^{-1}(x) = \overline{A}_n(x) \qquad (0 < x < 1).$$

We denote the corresponding density by $\overline{h}_n(y)$.

Lemma 3.9. *Under the above assumptions, $\{H_n(y)\}$ and $\{\overline{H}_n(y)\}$ are asymptotically indistinguishable.*

Proof. Let $\{Y_n\}$ be a sequence of random variables with c.d.f.'s $\{H_n(y)\}$. We must show that $h_n(Y_n)/\overline{h}_n(Y_n)$ converges stochastically to 1 as $n \to \infty$. For each n, let X_n be a random variable with distribution $H_n(\overline{A}_n(x))$, $0 < x < 1$. Then, the random variable $Y_n^* \equiv \overline{A}_n(X_n)$ has c.d.f. $H_n(x)$: thus we have the equivalent condition that $h_n(\overline{A}_n(X_n))/\overline{h}_n(\overline{A}_n(X_n))$ converges stochastically to 1 as $n \to \infty$.

Let $E_1(n)$ denote the event $\{p_n < X_n < 1-p_n\}$ and $E_2(n)$ denote its complement. It is easily shown that $P_{H_n\overline{A}_n}\{E_1(n)\}$ converges to 1 as $n \to \infty$. Given $E_1(n)$, $\overline{h}_n(\overline{A}_n(X_n)) = [\overline{B}_n'(X_n)]^{-1}$ and

$$\frac{h_n(\overline{A}_n(X_n))}{\overline{h}_n(\overline{A}_n(X_n))} = \frac{\phi(\overline{A}_n(X_n))}{t(X_n)}\left[1 - \frac{\varepsilon_n'(T(X_n))}{t(X_n)} - \frac{\varepsilon_n(T(X_n))T(X_n)}{t(X_n)}\right]\left[1 + \frac{\varepsilon_n'(\overline{A}_n(X_n))}{\phi(\overline{A}_n(X_n))}\right].$$

Assumptions (3.62) and (3.63) imply that the expressions in brackets converge to 1 as $n \to \infty$. Furthermore,

$$2\log\left(\frac{\phi(\overline{A}_n(X_n))}{t(X_n)}\right) = \tau^2(X_n) - \overline{A}_n^2(X_n) = 2\frac{\varepsilon_n(T(X_n))T(X_n)}{t(X_n)} - \left[\frac{\varepsilon_n(T(X_n))}{t(X_n)}\right]^2$$

when $E_1(n)$ occurs, and the latter expression converges to zero as $n \to \infty$. Thus, let $\{\varepsilon_m;\ m=1,2,\ldots\}$ be any null sequence. Then, the above arguments imply that for each m, there is some N_m such that if $n > N_m$,

$$P_{H_n \overline{A}_n}\left\{\left|1 - \frac{h_n(\overline{A}_n(X_n))}{\overline{h}_n(\overline{A}_n(X_n))}\right| < \varepsilon_m \;\middle|\; E_1(n)\right\} = 1.$$

Finally, let $\{\delta_m\}$ be another null sequence; let $\overline{N}_m$ by such that $n > \overline{N}_m$ implies $P_{H_n \overline{A}_n}\{E_1(n)\} > 1-\delta_m$. Then, $n \geqq \max\{N_o, N_m, \overline{N}_m\}$ implies

$$P_{H_n \overline{A}_n}\left\{\left|1 - \frac{h_n(\overline{A}_n(X_n))}{\overline{h}_n(\overline{A}_n(X_n))}\right| < \varepsilon_m\right\} > 1 - \delta_m.$$

We have shown that $h_n(\overline{A}_n(X_n))/\overline{h}_n(\overline{A}_n(X_n))$ converges stochastically to 1 as desired. QED

Thus, for all asymptotic probability calculations, $\overline{H}_n(x)$ can be used in place of $H_n(x)$. This will lead to more useful expressions from a computational standpoint.

Let $\overline{T}(n)$ denote the (K_n+1)-vector whose elements $\overline{T}_i(n)$ are given by

$$\overline{T}_i(n) = \begin{cases} q_n \overline{H}_n^{-1}\left(p_n + (i-1)\frac{L_n}{n}\right) & i=1, K_n+1 \\ \overline{H}_n^{-1}\left(p_n + (i-1)\frac{L_n}{n}\right) & i=2,3,\ldots,K_n \end{cases}$$

and let $\overline{t}(n)$ denote the (K_n+1)-vector with elements $\overline{t}_i(n) = \overline{h}_n(\overline{H}_n^{-1}(p_n+(i-1)L_n/n))$, $i=1,2,\ldots,K_n+1$. Also, let $\overline{e}(n)$ denote the (K_n+1)-vector with elements

$$\overline{e}_i(n) = \begin{cases} q_n \dfrac{\varepsilon_n(T_i(n))}{t_i(n)} & i=1,K_n+1 \\ \dfrac{\varepsilon_n(T_i(n))}{t_i(n)} & i=2,3,\ldots,K_n \end{cases}$$

Note that $\overline{T}(n) = \tilde{T}(n) - \overline{e}(n)$. Finally, let I_{K_n+1} denote the $(K_n+1)\text{x}(K_n+1)$ identity matrix, J_{K_n+1} denote the $(K_n+1)\text{x}(K_n+1)$ matrix all of whose elements are 1, and $M(n)$ denote the matrix $[I_{K_n+1} + (K_n+1)^{-1}J_{K_n+1}]$, with rows $m_i(n)$ and (i,j)th elements $m_{ij}(n)$, $i,j = 1,\ldots,K_n+1$.

Lemma 3.10. Under the alternatives (3.56) subject to (3.57) - (3.63),

$$(\frac{27}{2})^{1/2}[\frac{n^{5/2}p_n^{3/2}}{L_n^2}(\tilde{W}(n) - \tilde{\mu}(n))]$$

is asymptotically normally distributed with mean $-(27/2)^{1/2}d^*$ and variance 1.

Proof. Let $\overline{S}_1(n)$ denote the quantity

$$\frac{[\tilde{T}'(n)V^{-1}(n)\overline{T}(n)]^2}{[\tilde{T}'(n)V^{-1}(n)V^{-1}(n)\tilde{T}(n)][\overline{T}'(n)M(n)\overline{T}(n)]}$$

and let $\overline{S}_2(n) = \tilde{W}(n) - \overline{S}_1(n)$. Thus, $\overline{S}_1(n)$ is the constant term, and $\overline{S}_2(n)$ represents the remaining terms in a Taylor series expansion of

$\tilde{W}(n)$ about the point $\overline{T}(n)$. We first consider $\overline{S}_1(n)$. Denoting $S_1(n) - \overline{S}_1(n)$ by $\overline{\Delta}_n$, we have

$$\overline{\Delta}_n = \frac{[\gamma'(n)\tilde{T}(n)]^2}{\gamma'(n)\gamma(n)\tilde{T}'(n)\tilde{T}(n)} - \frac{[\gamma'(n)\overline{T}(n)]^2}{\gamma'(n)\gamma(n)\overline{T}'(n)M(n)\overline{T}(n)} .$$

Recalling that $\gamma'(n) = 2(L_n-1)\tilde{T}'(n) + e'(n)$, $\tilde{T}'(n)\tilde{T}(n) = (n/L_n)(1 + O(p_n\log(1/p_n)))$ and $\gamma'(n)\gamma(n) = 4(L_n-1)^2(n/L_n)(1+o(1))$, we have

$$\gamma'(n)\gamma(n)\tilde{T}'(n)\tilde{T}(n)\overline{T}'(n)M(n)\overline{T}(n) = 4(L_n-1)^2(n/L_n)^3(1+o(1)),$$

and, neglecting smaller order terms, $\overline{\Delta}_n$ can be written as

$$\frac{L_n}{n}\left\{\begin{array}{l} \overline{e}'(n)\overline{e}(n) + L_n^{-1}\overline{e}'(n)e(n) + (2nL_n)^{-1}[e'(n)\tilde{T}(n)][\overline{e}'(n)e(n)] \\ + (2n)^{-2}[\overline{e}'(n)\overline{e}(n)][e'(n)\tilde{T}(n)]^2 + n^{-1}[e'(n)\tilde{T}(n)][\overline{e}'(n)\overline{e}(n)] \end{array}\right\}$$

$$- \frac{L_n}{n^2}\left\{\begin{array}{l} [e'(n)\tilde{T}(n)][\overline{e}'(n)\tilde{T}(n)] + (2n)^{-1}[e'(n)\tilde{T}(n)]^2[e'(n)\tilde{T}(n)] \\ L_n^{-1}[\overline{e}'(n)\tilde{T}(n)]^2 + [\overline{e}'(n)\tilde{T}(n)][\overline{e}'(n)e(n)] + (4L_n)^{-1}[\overline{e}'(n)e(n)]^2 \end{array}\right\}$$

$$- \frac{L_n^2}{n^2}[\sum_{i=1}^{K_n+1}\overline{T}_i(n)]^2\{1 + n^{-1}e'(n)\tilde{T}(n) + (2n)^{-2}[e'(n)\tilde{T}(n)]^2\}.$$

Finally, using the definitions of $e(n)$ and $\overline{e}(n)$, approximating sums by integrals, and neglecting smaller order terms, it is easily shown that $\overline{\Delta}_n = \rho(n,\varepsilon_n)$.

We next consider $\overline{S}_2(n)$. To do so, we return to the notation of lemma

3.7. We will first show that

$$\tilde{w}^{(k)}_{i_1 i_2 \ldots i_k}(\overline{T}(n)) = \tilde{w}^{(k)}_{i_1 i_2 \ldots i_k}(\tilde{T}(n)) \quad (1+o(1)) \tag{3.64}$$

as $n \to \infty$ for all $k=1,2,\ldots$ and all indices $i_k \in \{1,2,\ldots,K_n+1\}$. By (3.45), (3.64) follows immediately from the following relations:

$$\frac{A_i'(\overline{T}(n))}{A(\overline{T}(n))} = \frac{A_i'(\tilde{T}(n))}{A(\tilde{T}(n))}(1+o(1)) \qquad \frac{B_i'(\overline{T}(n))}{B(\overline{T}(n))} = \frac{B_i'(\tilde{T}(n))}{B(\tilde{T}(n))}(1+o(1))$$

$$\frac{A_{ij}''(\overline{T}(n))}{A(\overline{T}(n))} = \frac{A_{ij}''(\tilde{T}(n))}{A(\tilde{T}(n))}(1+o(1)) \qquad \frac{B_{ij}''(\overline{T}(n))}{B(\overline{T}(n))} = \frac{B_{ij}''(\tilde{T}(n))}{B(\tilde{T}(n))}(1+o(1)) \tag{3.65}$$

as $n \to \infty$ for $i,j = 1,2,\ldots,K_n+1$.

To prove (3.65), we first consider the quantity $\gamma'(n)\overline{e}(n)$, which can be written for asymptotic purposes as

$$2n\int_{p_n}^{1-p_n} \frac{T(x)}{t(x)}\varepsilon_n(T(x))dx + \frac{L_n^2}{3n}\int_{p_n}^{1-p_n} \frac{T^3(x)}{t^3(x)}\varepsilon_n(T(x))dx.$$

Applying the mean value theorem for integrals (e.g. Bartle [1,p.327]), (3.62), (C.8) and (C.10), we have $\gamma'(n)\overline{e}(n) = o(nb^{-1}(n)(\log 1/p_n)^{1/2})$ as $n \to \infty$. Using this fact,

$$\frac{A_i'(\overline{T}(n))}{A(\overline{T}(n))} = \frac{2\gamma_i(n)\gamma'(n)\overline{T}(n)}{[\gamma'(n)\overline{T}(n)]^2}$$

$$= \left[\frac{A_i'(\tilde{T}(n))}{A(\tilde{T}(n))} - \frac{\gamma_i(n)\gamma'(n)\overline{e}(n)}{[\gamma'(n)\tilde{T}(n)]^2}\right]\left[\frac{\gamma'(n)\tilde{T}(n)}{\gamma'(n)\overline{T}(n)}\right]^2$$

$$= [\frac{A_i(\tilde{T}(n))}{A(\tilde{T}(n))} + o(\frac{L_n \log(1/p_n)}{nb(n)})][1 + o(b^{-1}(n)(\log 1/p_n)^{1/2})]$$

$$= \frac{A_i'(\tilde{T}(n))}{A(\tilde{T}(n))}[1 + o(b^{-1}(n)(\log 1/p_n)^{1/2})],$$

where the last equality follows from (3.44). The remaining relations in (3.65) are similarly established and (3.64) follows. Next we note that since

$$\bar{t}_i(n) = t_i(n)[1 - \frac{\varepsilon_n(T_i(n))T_i(n)}{t_i(n)} - \frac{\varepsilon_n'(T_i(n))}{t_i(n)}] = t_i(n)[1 + \delta(i,n)]$$

where $\delta(i,n)$ converges to zero as $n \to \infty$ uniformly for $1 \leq i \leq K_n+1$, the covariance matrix of $(\tilde{Y}_1,\ldots,\tilde{Y}_{K_n+1})$ under the sequence of alternatives $\{\bar{H}_n(x)\}$ may be taken to be $V(n)$ as defined in (3.6) for all asymptotic purposes. Thus, we have shown that

$$E\{\bar{S}_2(n)\} = E\{S_2(n) + S_3(n)\}(1+o(1)) = o_p(\tilde{\sigma}(n)),$$

and

$$\sigma\{\bar{S}_2(n)\} = \sigma\{S_2(n) + S_3(n)\}(1+o(1)) = \tilde{\sigma}(n)(1+o(n)).$$

Thus, under the sequence of alternatives $\{\bar{H}_n(x)\}$, the quantity

$$\tilde{\sigma}^{-1}(n)[\tilde{W}(n) - \bar{S}_1(n)] = \tilde{\sigma}^{-1}(n)[\tilde{W}(n) - \tilde{\mu}(n)] + \tilde{\sigma}^{-1}(n)\bar{\Delta}_n$$

is asymptotically distributed as standard normal. Applying (3.61) and lemma 3.9, the proof of lemma 3.10 is complete. QED

Theorem 3.4 follows immediately from lemma 3.10.

3.5 Extensions

The results of section 3.4 are easily extended to the statistic $W^*(n)$, as summarized in

Corollary 3.1. Assume (3.56) through (3.60) hold as well as

(i) (3.62 and (3.63) hold with $C(n)$ replaced by $C^*(n) = \frac{L_n^2}{n^{3/2} p_n^{3/2}} C(n)$

(ii) $\lim_{n\to\infty} n\rho^*(n,\varepsilon_n) = d_1^*, \quad 0 < d_1^* < \infty$

where

$$\rho^*(n,\varepsilon_n) = \int_{p_n}^{1-p_n} \left[\frac{\varepsilon_n(T(x))}{t(x)}\right]^2 dx - \left\{ \int_{p_n}^{1-p_n} \frac{\varepsilon_n(T(x))}{t(x)} dx\right\}^2$$

$$- \left\{ \int_{p_n}^{1-p_n} \frac{T(x)}{t(x)} \varepsilon_n(T(x)) dx \right\}^2 .$$

Then, $n(1 - W^*(n)) - a_n^*$ converges in law to $\Psi(y - d_1^*)$ as $n \to \infty$.

Proof. Using lemma 3.9 (with $q_n \equiv 1$) and expanding $W^*(n)$ in a Taylor series about $T(n)$, the constant term in the expansion is given by $1 - \rho^*(n,\varepsilon_n)$ as $n \to \infty$. The methods used in the ɔof of lemma 3.10 apply immediately to the quadratic and higher order terms, yielding the fact that the differences between these terms in the expansions under H_o and H_1 are $o_p(1/n)$. To see that the linear term in the expansion is also $o_p(1/n)$, we use (3.24) (with $\gamma(n)$ replaced by $T(n)$) together with assumption (i) above to obtain the result that $\tilde{W}_i^{(1)}(\overline{T}(n))$ is $O(L_n(\log 1/p_n)^{3/2}/nb(n))$ as $n \to \infty$. Hence, using (3.16), the variance of the linear term is of smaller order than $(\log 1/p_n)^3/nb^2(n)$, which is $O(1/n^2)$. Applying Theorem 3.2, the proof of the corollary is complete. QED

We remark that the Schwarz inequality implies that the "metric" $\rho^*(n,\varepsilon_n)$ is always positive. The same cannot always be said to hold for $\rho(n,\varepsilon_n)$. In fact, the $\tilde{W}$-test can have very interesting power functions, as illustrated in example 3.2 below. The next result, however, gives one example of a sufficient condition for $\rho(n,\varepsilon_n) = \rho^*(n,\varepsilon_n)$.

Corollary 3.2. *Assume that (3.56) through (3.63) hold as well as*

(A) $\lim\limits_{n\to\infty} L_n p_n^2/n^{3/4} = 0$

(B) $\dfrac{\varepsilon_n(T(x))}{t(x)}$ *is of the form* $\sum\limits_{m=0}^{\infty} c_m(n)T^m(x)$

where the constants $\{c_m(n)\}$ *satisfy*

$$c_m(n) = O(c_{m-1}(n)p_n) \quad m=1,2,\ldots$$

Then,

$$\rho(n,\varepsilon_n) = \rho^*(n,\varepsilon_n) = 2c_0^2(n)p_n(1+o(1))$$

as $n\to\infty$.

Proof. Let

$$\rho^{(1)} = \frac{1}{3}\frac{L_n^2}{n^2}\sum_{m=0}^{\infty} c_m(n)\Big[\int_{p_n}^{1-p_n}\frac{T^{3+m}(x)}{t^2(x)}\,dx\int_{p_n}^{1-p_n}T^2(x)dx$$

$$-\int_{p_n}^{1-p_n}\frac{T^4(x)}{t^2(x)}\,dx\int_{p_n}^{1-p_n}T^{1+m}(x)dx\Big]$$

$$= \frac{1}{3}\frac{L_n^2}{n^2}\sum_{m=1}^{\infty} c_{2m+1}(n)\frac{(\log 1/p_n)^{m+1}}{p_n}(1 + O(p_n\log(1/p_n)))$$

$$= O\left(\frac{L_n^2}{n^2}c_0(n)\;(p_n\log 1/p_n)^2\right)$$

since the expression in brackets is identically zero for $m = 0,1,2$ and all even integers thereafter. Note that

$$\int_{p_n}^{1-p_n} \left[\frac{\varepsilon_n(T(x))}{t(x)}\right]^2 dx = c_0^2(n)(1-2p_n+0(p_n^2))$$

$$\left\{\int_{p_n}^{1-p_n} \frac{\varepsilon_n(T(x))}{t(x)} dx\right\}^2 = c_0^2(n)((1-2p_n)^2+0(p_n^2))$$

$$\left\{\int_{p_n}^{1-p_n} \frac{T(x)}{t(x)}\varepsilon_n(T(x))dx\right\}^2 = c_1^2(n)(1+0(p_n^2)) = c_0^2(n)0(p_n^2).$$

Combining the above and neglecting smaller order terms, we see that $\rho(n,\varepsilon_n) = 2c_0^2 p_n + \rho^{(1)}$, and $\rho^*(n,\varepsilon_n) = c_0^2 p_n$. Finally, assume that

$$\lim_{n\to\infty} \frac{n^{5/2}p_n^{3/2}}{L_n^2} (c_0^2(n)p_n) = \frac{d^*}{2} .$$

This implies that $c_0(n) = 0(L_n/(np_n)^{5/4}$, so that

$$\frac{n^{5/2}p_n^{3/2}}{L_n^2} \rho^{(1)} = 0\left(\frac{L_n p_n^{9/4}(\log 1/p_n)^2}{n^{3/4}}\right) = o(1)$$

by assumption (A). Thus, under (3.61), $\rho^{(1)}$ is of smaller order than $c_0^2(n)p_n$, and $\rho(n,\varepsilon_n) = \rho^*(n,\varepsilon_n) = 2c_0^2(n)p_n(1+o(1))$ as desired. QED

It is clear from the form of the power function of the $\tilde{W}$-test that the rate of approach of alternatives to the null hypothesis which yields a nontrivial power increases as the order of L_n decreases. Hence, assymption (A) is quite reasonable. Examples 3.1 and 3.2 below illustrate assumption (B) when (A) does and does not hold, respectively.

Before turning to the examples, we introduce two additional statistics motivated by the results above. In both $\rho(n,\varepsilon_n)$ and $\rho^*(n,\varepsilon_n)$, the term

$$-\left\{\int_{p_n}^{1-p_n} \frac{\varepsilon_n(T(x))}{t(x)}\,dx\right\}^2$$

arose because $\tilde{D}^2(n)$ was used in the demoninators of $\tilde{W}(n)$ and $\tilde{W}^*(n)$. In particular, we found that replacing $\tilde{D}^2(n)$ by $\tilde{Y}'\tilde{Y}$ results in the elimination of precisely the above term from each of the two metrics. That, of course, does not guarantee that the tests based on the modified statistics will have improved power, since the distributions of the statistics under the null hypothesis, and hence the critical regions of the tests, would be expected to change also. However, it turns out that the asymptotic distribution of the modified $\tilde{W}(n)$ does <u>not</u> change, and the distribution of the modified $W^*(n)$ changes only slightly.

Formally, we define the following two statistics:

$$\tilde{\tilde{W}}(n) = \frac{[\tilde{T}'(n)V^{-1}(n)\tilde{Y}]^2}{[\tilde{T}'(n)V^{-1}(n)V^{-1}(n)\tilde{T}(n)]\tilde{Y}'\tilde{Y}}$$

$$\tilde{\tilde{W}}^*(n) = \frac{[T'(n)\tilde{Y}]^2}{[T'(n)T(n)]\tilde{Y}'\tilde{Y}} \quad .$$

The asymptotic distributions of these statistics under the null and alternative hypotheses are summarized in the last two corollaries of this section.

Corollary 3.3. Under H_o,

(i) $(\frac{27}{2})^{1/2} \frac{n^{5/2} p_n^{3/2}}{L_n^2} (\tilde{\tilde{W}}(n) - \tilde{\mu}(n))$ converges in law to $\Phi(x)$

(ii) $n(1 - \tilde{\tilde{W}}^*(n)) - a_n^{**}$ converges in law to $\Psi_1(y)$, where

$$a_n^{**} = \frac{L_n}{n} \sum_{i=1}^{K_n+1} \sum_{j=1}^{K_n+1} [\frac{L_n}{n} T_i(n) T_j(n) - \delta_{ij}] v_{ij}(n)$$

and $\Psi_1(y)$ is the distribution of

$$\sum_{k=2}^{\infty} \frac{(X_k^2 - 1)}{k}$$

with $X_2, X_3, \ldots$ i.i.d. standard normal random variables.

Proof. (i) Consider the effect of replacing $\tilde{D}^2(n)$ by $\tilde{Y}'\tilde{Y}$ on the expansion (3.19). Since $\tilde{T}. = 0$, $S_1(n)$ is unchanged. Similarly, $\tilde{w}_i^{(1)}(\tilde{T}(n))$ is unchanged, so that $S_2(n)$ remains the same. The only effect on $\tilde{w}_{ij}^{(2)}(\tilde{T}(n))$ is to eliminate the term $(K_n+1)^{-1}$, but this has no bearing on the asymptotic order of $S_3(n)$. In fact, (3.44) still holds (except that $B''_{ij}(.) = 0$ for $i \neq j$) so that the effect on $S_4(n)$ is asymptotically negligible. In summary, the asymptotic distribution of $\tilde{\tilde{W}}(n)$ is the same as that of $\tilde{W}(n)$ under the null hypothesis.

(ii) The proof of this result appears at the end of Appendix B. QED

Corollary 3.4. Assume (3.56) through (3.60) hold.

(i) Assume (3.61) through (3.63) hold with $\rho(n,\varepsilon_n)$ replaced by

$$\rho_1(n,\varepsilon_n) = \rho(n,\varepsilon_n) + \{ \int_{p_n}^{1-p_n} \frac{\varepsilon_n(T(x))}{t(x)} dx \}^2 .$$

Then,

$$\left(\frac{27}{2}\right)^{1/2} \frac{n^{5/2} p_n^{3/2}}{L_n^2} (\tilde{\tilde{W}}(n) - \tilde{\mu}(n))$$

converges in law to $\Phi(x + (27/2)^{1/2} d^*)$ as $n \to \infty$.

(ii) Assume (i) and (ii) of Corollary 3.1 hold with $\rho^*(n,\varepsilon_n)$ replaced by

$$\rho_1^*(n,\varepsilon_n) = \rho^*(n,\varepsilon_n) + \left\{ \int_{p_n}^{1-p_n} \frac{\varepsilon_n(T(x))}{t(x)} dx \right\}^2 .$$

Then, $n(1 - \tilde{\tilde{W}}^*(n)) - a_n^{**}$ converges in law to $\Psi_1(y-d_1^*)$ as $n \to \infty$.

Proof. (i) The change in the metric from ρ to ρ_1 is easily verfied. Since (3.64) remains valid, the remainder of the proof of lemma 3.10 applies and the result follows. (ii) The change in the metric from ρ^* to ρ_1^* is again easily verified. The proof of corollary 3.1 applies to the linear and higher order terms in the expansion of $\tilde{\tilde{W}}^*(n)$ and the result follows. QED

Thus, the power of the test based on $\tilde{\tilde{W}}(n)$ is strictly greater than the power of the test based on $\tilde{W}(n)$. We cannot make such a strong statement about the power of the tests based on $\tilde{W}^*(n)$ and $\tilde{\tilde{W}}^*(n)$, but, in general, $\tilde{\tilde{W}}^*(n)$-tests are more powerful and often hyper-efficient with respect to $\tilde{W}^*(n)$-tests.

3.6 Examples

Example 3.1. Let $H_n(x) = (1-\delta_n)\Phi(x) + \delta_n\Phi(x+\lambda_n)$, where $0 < \delta_n < 1$, and take $L_n = n^{3/4}(\log 1/p_n)^{-1}$. From assumption (3.4), this is "almost"

as small as L_n can be made. Finally, let $\lambda_n \equiv p_n$. Then

$$\varepsilon_n(x) = \delta_n(\Phi(x+p_n) - \Phi(x)),$$

or

$$\varepsilon_n(\Phi^{-1}(y)) = \delta_n(\Phi(\Phi^{-1}(y)+p_n) - y).$$

Using a Taylor series expansion and rearranging terms,

$$\Phi(\Phi^{-1}(y)+p_n) = y + p_n\phi(\Phi^{-1}(y)) \left\{ \begin{array}{l} 1 - \frac{p_n^2}{6} - (\frac{p_n}{2} - \frac{p_n^3}{8})\Phi^{-1}(y) + \frac{p_n^2}{6}(\Phi^{-1}(y))^2 \\ \\ - \frac{p_n^3}{24}(\Phi^{-1}(y))^3 + 0(p_n^4(\log 1/p_n)^4) \end{array} \right\},$$

so that

$$\frac{\varepsilon_n(\Phi^{-1}(y))}{\phi(\Phi^{-1}(y))} = p_n\delta_n \left\{ \begin{array}{l} 1 - \frac{p_n^2}{6} - (\frac{p_n}{2} - \frac{p_n^3}{8})\Phi^{-1}(y) + \frac{p_n^2}{6}(\Phi^{-1}(y))^2 \\ \\ - \frac{p_n^3}{24}(\Phi^{-1}(y))^3 + 0(p_n^4(\log 1/p_n)^4) \end{array} \right\}.$$

Thus, both (A) and (B) of corollary 3.2 are satisfied, and

$$\rho(n,\varepsilon_n) = \rho^*(n,\varepsilon_n) = 2p_n(p_n\delta_n)^2(1+o(1)).$$

Also,

$$\rho_1(n,\varepsilon_n) = \rho_1^*(n,\varepsilon_n) = (p_n\delta_n)^2(1+o(1)).$$

Finally, for purposes of comparison, the Kolmogorov-Smirnov distance is given by

$$\sup_{-\infty<x<\infty} |H_n(x) - \Phi(x)| = \sup_{-\infty<x<\infty} |\varepsilon_n(x)| = (2\pi)^{-1/2} p_n \delta_n .$$

The conditions yielding non-trivial powers and the powers themselves for the various tests we have considered are summarized in Table 3.2. Let "$A_n << B_n$" signify that the sequence of tests based on the statistics $\{A_n\}$ in hyper-efficient with respect to the sequence of tests based on the statistics $\{B_n\}$. Then, for the sequence of alternatives considered in this example,

$$T(n;\hat{\theta}_1,\hat{\theta}_2) << \tilde{W}(n) << \tilde{\tilde{W}}(n) << \tilde{W}^*(n) << \begin{cases} \tilde{\tilde{W}}^*(n) \\ D \end{cases},$$

where D denotes the two-sided Kolmogorov-Smirnov statistic.

Example 3.2. We again let $H_n(x) = (1-\delta_n)\Phi(x) + \delta_n\Phi(x+p_n)$, except this time we choose

$$L_n = n^{.8} \qquad p_n = n^{-(\log n)^{-2/5}} .$$

Feasibility conditions (3.2), (3.3), and (3.4) are easily verified for these functions. Using results in the proof of corollary 3.2, we have, neglecting smaller order terms,

$$\rho(n,\varepsilon_n) = -\frac{1}{72}((L_n/n)(\log 1/p_n))^2 p_n\delta_n + 2p_n\delta_n^2 .$$

Note that

$$p_n\delta_n^2 = o((L_n/n)^2(\log 1/p_n)^2 p_n\delta_n) \text{ iff } \delta_n = o((L_n/n)^2(\log 1/p_n)^2),$$

and

TABLE 3.2

ASYMPTOTIC POWERS UNDER $H_n(x) = (1-\delta_n)\Phi(x) + \delta_n\Phi(x+p_n)$

Sequence of tests based on	Condition yielding non-trivial power	Power of level-α test[†]
$\tilde{W}(n)$	$n^{1/2}p_n^{9/4}(\log 1/p_n)\delta_n = d_1$	$\Phi(c_{\Phi,\alpha} + (54)^{1/2}d_1^2)$
$\tilde{\tilde{W}}(n)$	$n^{1/2}p_n^{7/4}(\log 1/p_n)\delta_n = d_2$	$\Phi(c_{\Phi,\alpha} + (27/2)^{1/2}d_2^2)$
$\tilde{W}^*(n)$	$n^{1/2}p_n^{3/2}\delta_n = d_3$	$1 - \Psi(c_{\Psi,1-\alpha} - d_3)$
$\tilde{\tilde{W}}^*(n)$	$n^{1/2}p_n\delta_n = d_4$	$1 - \Psi_1(c_{\Psi_1,1-\alpha} - d_4)$
D	$n^{1/2}p_n\delta_n = d_5$	††
$T(n;\hat{\theta}_1,\hat{\theta}_2)$	$L_n^{1/4}n^{1/4}p_n\delta_n = d_6$	$\Phi(c_{\Phi,\alpha} + 1.99d_6^2)$

†where $c_{F,\alpha}$ is defined by $\int_{-\infty}^{c_{F,\alpha}} dF(x) = \alpha$

††closed form expression does not exist

$$\lim_{n\to\infty} \frac{n^{5/2}p_n^{3/2}}{L_n^2} \frac{L_n^2(\log 1/p_n)^2}{n^2} p_n\delta_n < \infty$$

implies that

$$\delta_n = O(n^{-1/2}p_n^{-5/2}(\log 1/p_n)^{-2}) = o(\frac{L_n^2(\log 1/p_n)^2}{n^2}) .$$

In Figure 3.3, we use the above facts in conjection with assumption (3.61) to plot the power of the $\tilde{W}$-test as a function of the rate at which δ_n approaches zero. When $\delta_n = \text{const}\cdot n^{-1/2}p_n^{-5/2}(\log 1/p_n)^{-2}$, $\rho(n,\varepsilon_n) < 0$, and any power between 0 and α is obtainable by choosing the constant appropriately. For $\delta_n = \text{const}\cdot(L_n/n)^2(\log 1/p_n)^2$, any power between 0 and 1 is obtainable by choosing the constant to solve the appropriate quadratic equation.

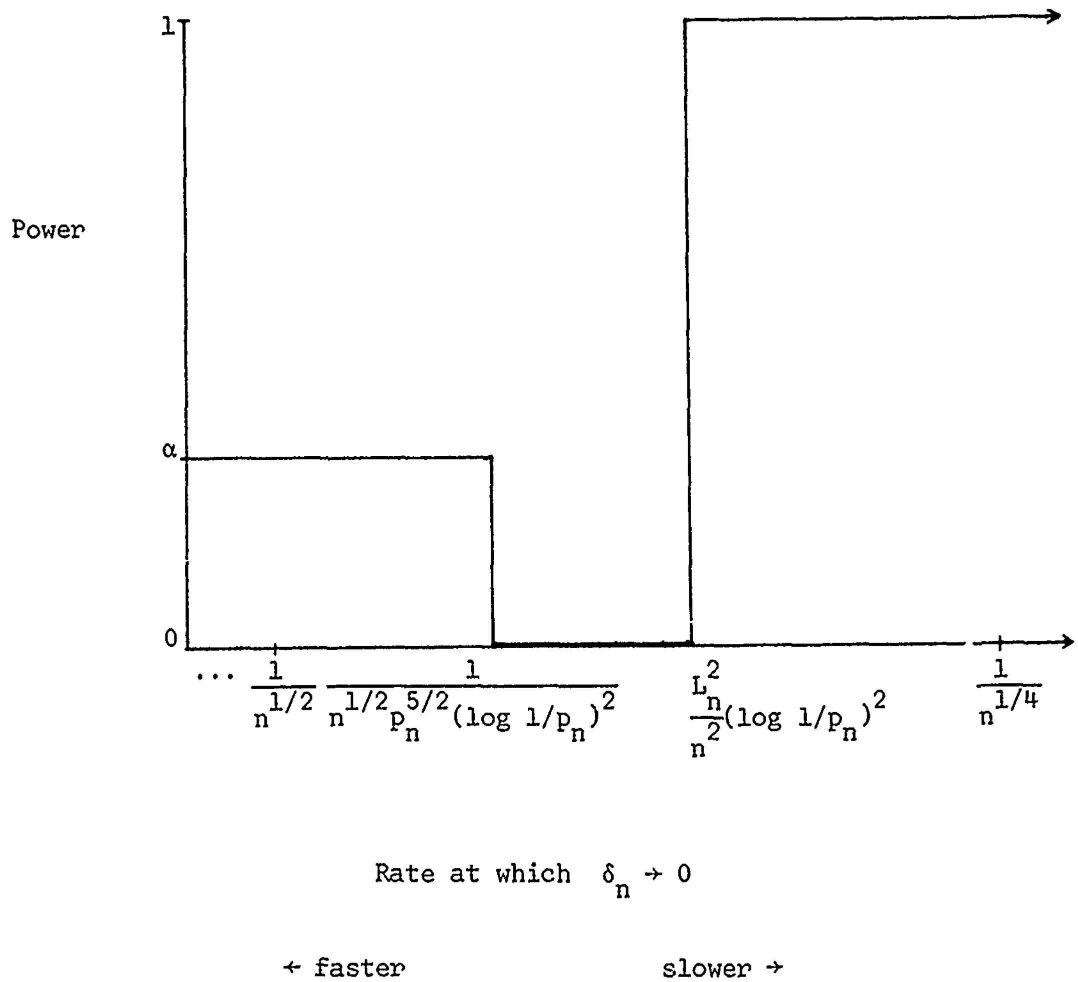
Figure 3.3 illustrates a rather interesting phenomenon--a test which is both biased against a sequence of contiguous alternatives and consistent. As alternatives approach the null hypothesis at a very fast rate, the test cannot discriminate between the alternatives and the hypothesized distribution, and the power is equal to the level of the test. As the rate of approach slows down, the test can discriminate, but because of the bias, the power <u>decreases</u> rather than increases. Eventually, if the rate of approach slows down enough, the consistency property of the test takes over and the power increases to one. What is most interesting is that the rates at which the bias sets in and the consistency sets in are <u>different</u>, resulting in a set of alternatives which are never detected by the test.

The author knows of no analogs to the above behavior in the literature.

FIGURE 3.3

ASYMPTOTIC POWER OF THE $\tilde{W}$-TEST

AS A FUNCTION OF THE RATE AT WHICH δ_n APPROCHES ZERO



Bias has been exhibited for both the Kolmogorov-Smirnov test (Massey [18]) and the Cramér-von Mises test (Thompson [36]). However, in each case, the bias was proved only for a fixed alternative and fixed n finite; to our knowledge, analysis similar to the above in which contiguous alternatives are considered has not been reported.

3.7 Concluding Remarks

There are at least three results above which deserve further emphasis in that they suggest possible areas for future research.

First, the large-sample results above for the tests based on a gradually increasing number of order statistics stand in contrast to the empirical, small-sample, results obtained for the analogous statistics based on all n order statistics. On the basis of the empirical studies cited in section 3.1, the consensus in the literature has been that in using the Shapiro-Francia test, an approximation to the Wilk-Shapiro test, one gains computationally but pays at least a small price in terms of sensitivity to alternative distributions. Above, we found that it was $\tilde{W}$ that approximated $\tilde{W}^*$, and not vice versa, in the sense that the errors in the approximation controlled the asymptotic distribution of $\tilde{W}$ and caused the associated test to be inferior to the $\tilde{W}^*$-test in discriminating contiguous alternatives. In section 3.2.4, we mentioned a possible approach toward resolving the open question of the relation between the large-sample and small-sample results.

Second, the results for the modified statistics $\tilde{\tilde{W}}$ and $\tilde{\tilde{W}}^*$ suggest that modifying W and W^* in the same manner would lead to superior tests. It would be useful to carry out empirical, small-sample, studies on such tests. In view of the widespread acceptance of the W- and

and W^*-tests, it would be of considerable practical value if such studies should yield results which agree with the large-sample results obtained here.

Finally, the test based on $\tilde{\tilde{W}}^*$ is worthy of further study in its own right. In example 3.1 we found that even against alternatives which are particularly easily discriminated by the Kolmogorov-Smirnov test (i.e. the Kolmogorov-Smirnov distance is maximized at the origin), the $\tilde{\tilde{W}}$-test performed well. Since the metric $\rho_1^*(n,\varepsilon_n)$ heavily weights disturbances in the tails, it should not be difficult to find situations where the $\tilde{\tilde{W}}^*$-test is very efficient relative to the Kolmogorov-Smirnov test, and perhaps to the other EDF tests as well. It would also be interesting to study the empirical behavior of the $\tilde{\tilde{W}}^*$-test, but, as pointed out in section 3.2.4, for moderate sample sizes the asymptotic theory is inadequate, and even the distribution of $W^*(n)$ under the null hypothesis would have to be computed empirically.

APPENDIX A

MONTE CARLO RESULTS FOR THE WEISS STATISTIC UNDER H_0

In this appendix we describe the results of a Monte Carlo study of the distributions of the statistics $T(n;\theta_1,\theta_2)$ and $T(n;\hat{\theta}_1,\hat{\theta}_2)$ (defined in section 2.2.2) under the null hypothesis of normality.

For each of the sample sizes $n = 100$, $n = 200$, and $n = 300$, we consider three choices of the functions L_n, K_n, and p_n. These choices are labeled cases I, II, III as defined below.

Case I: L_n = greatest integer $\leq n^{1-\delta_n}$ where $\delta_n = (\log n)^{-1/4}$

K_n = greatest even integer $\leq n^{\delta_n}[1 - 2n^{-\delta_n^2}]$

$p_n = (n - K_n L_n)/2n$.

Case II: L_n = greatest even integer $\leq n^{\delta}$ where $\delta = .7$

p_n taken as close as possible to the value in case I such that

np_n is an integer and $K_n \geq 2$

$K_n = (n - 2np_n)/L_n$.

Case III: same as case II with $\delta = .8$.

Since we were particularly interested in studying different choices of L_n, the values of p_n were made as uniform as possible in order to minimize the effect of moving into the tails at different rates. The values of the functions for the different cases and sample sizes are summarized in Table A.1.

TABLE A.1

PARAMETER SETTINGS FOR MONTE CARLO STUDY

	n								
	100			200			300		
	Case			Case			Case		
Parameter	I	II	III	I	II	III	I	II	III
L_n	4	24	39	6	40	69	7	54	95
K_n	16	3	2	26	4	2	32	4	2
p_n	.18	.14	.11	.11	.10	.115	.126	.14	.18

For each case and value of n, 100 sets of n $N(2,4)$ random variables were generated using the Box-Mueller method, and the statistics $T(n;\theta_1,\theta_2)$ and $T(n;\hat{\theta}_1,\hat{\theta}_2)$ were computed. The ordered values of the statistics are plotted on normal probability paper in figures A.1 through A.6. Each figure contains the results of cases I, II, and III for a specified value of n, the odd numbered figures when θ_1 and θ_2 were specified, the even numbered figures when they were estimated from the samples. Deviations from linearity in the figures correspond to deviations from normality of the empirical distributions of the statistics.

The study was admittedly limited in scope; those conclusions we felt were valid are discussed in section 2.4.1.

FIGURE A.1

NORMAL PLOT OF ORDERED $T(100;\theta_1,\theta_2)$ VALUES FOR CASES I(⊙),II(⊙),III(●)

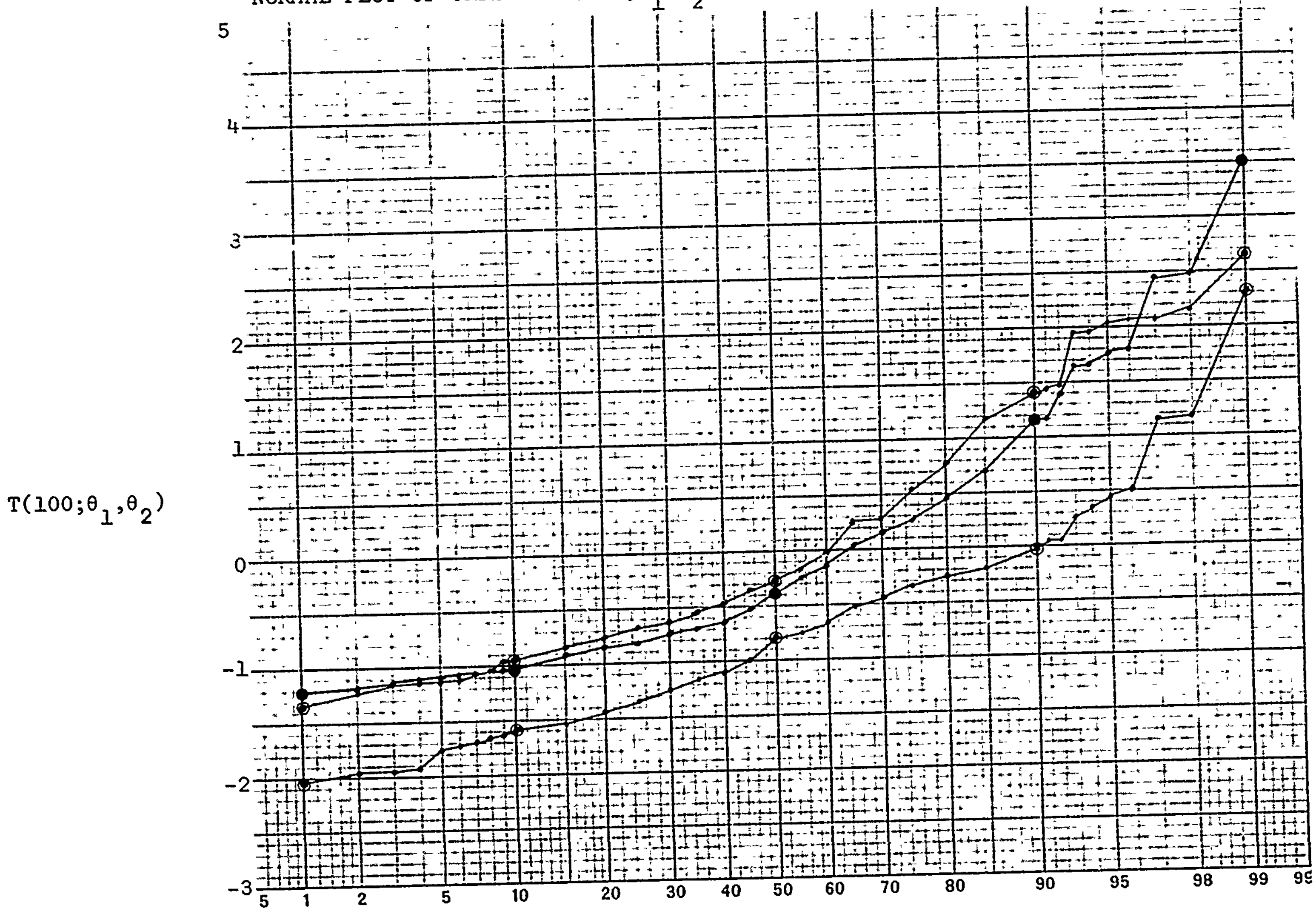


FIGURE A.2
NORMAL PLOT OF ORDERED $T(100;\hat{\theta}_1,\hat{\theta}_2)$ VALUES FOR CASES I(⊚),II(⊛),III(●)

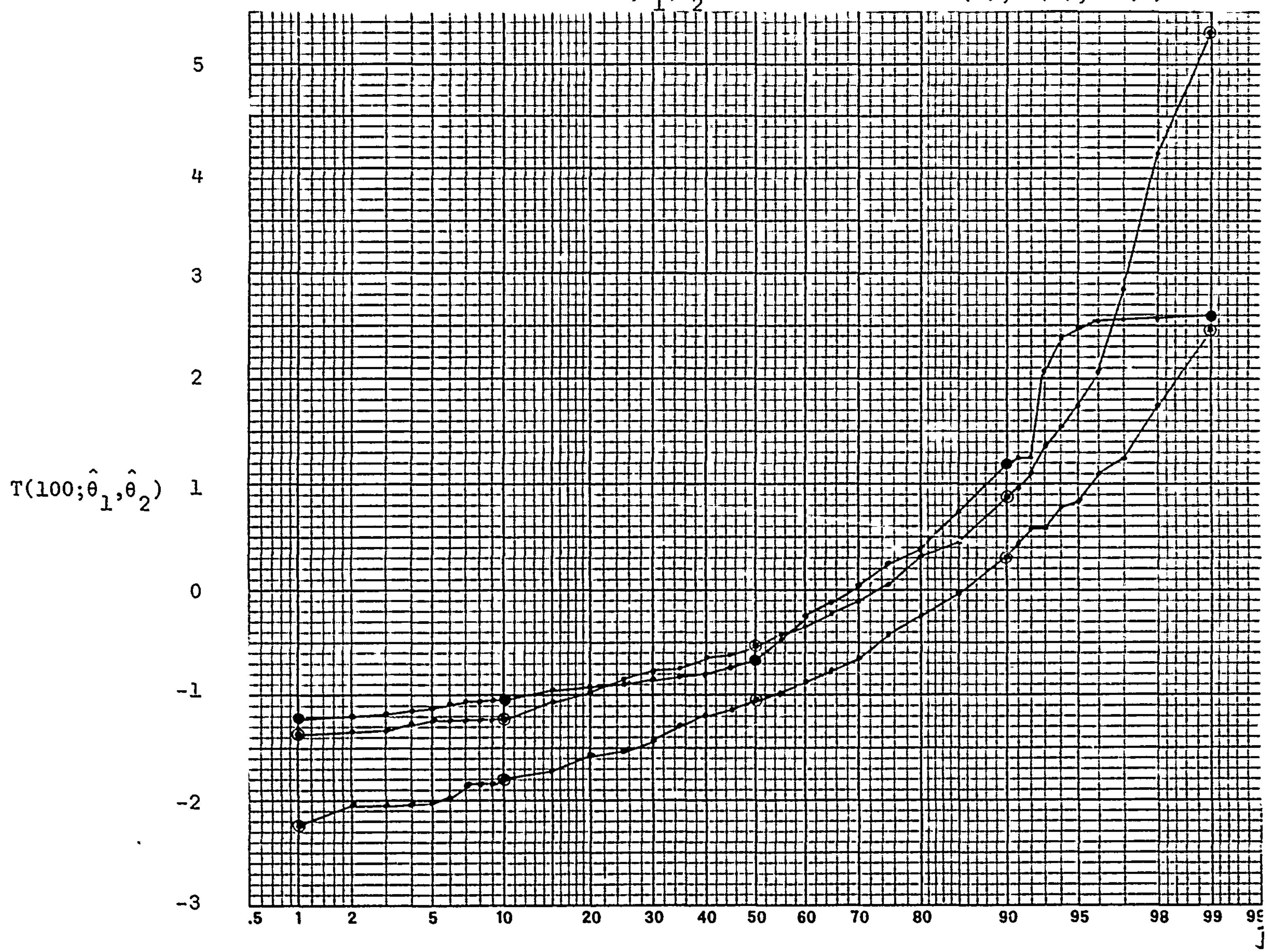


FIGURE A.3
NORMAL PLOT OF ORDERED $T(200;\theta_1,\theta_2)$ VALUES FOR CASES I(⊙),II(⊙),III(●)

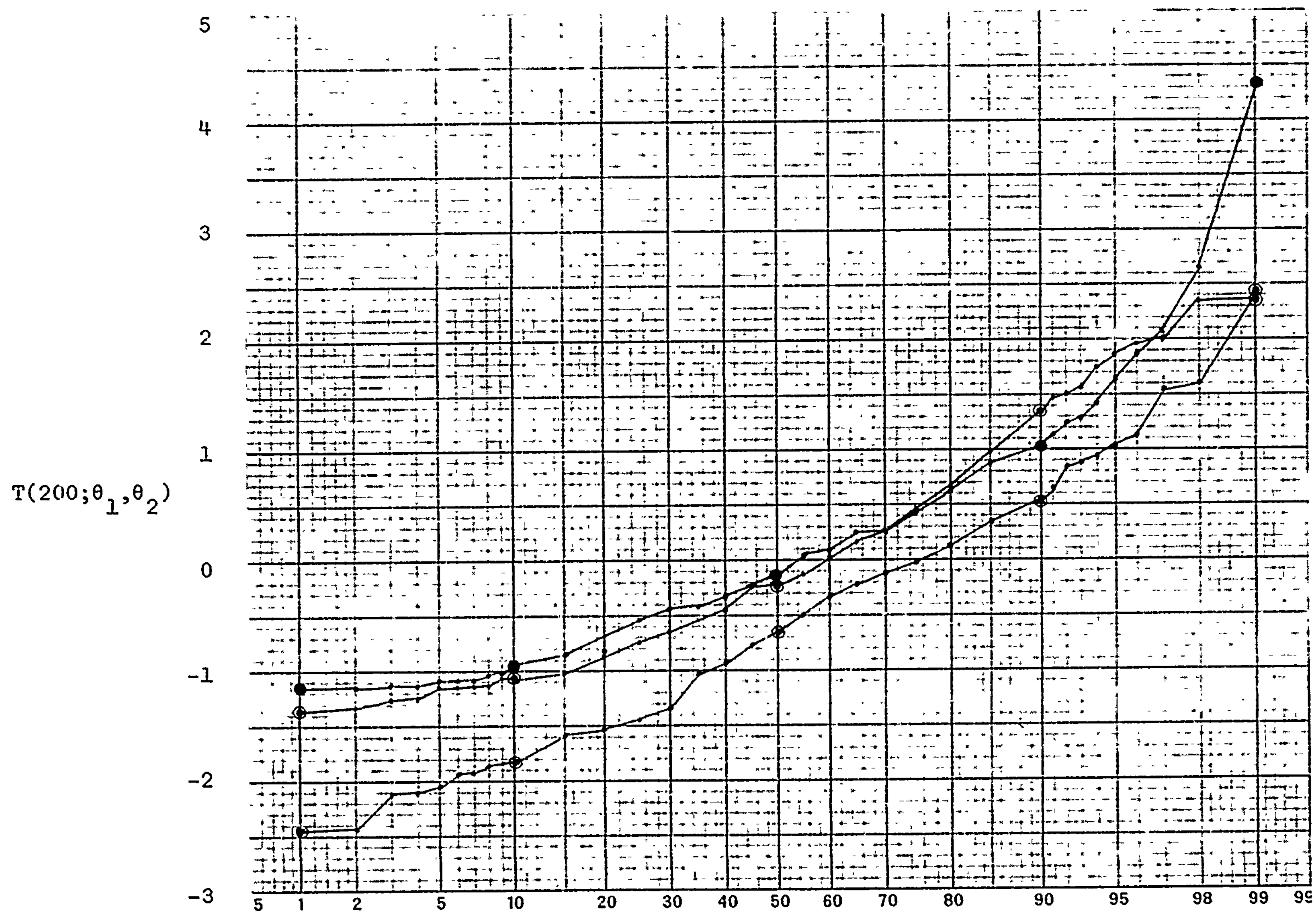


FIGURE A.4

NORMAL PLOT OF ORDERED $T(200;\hat{\theta}_1,\hat{\theta}_2)$ VALUES FOR CASES I(⊛), II(⊛), III(●)

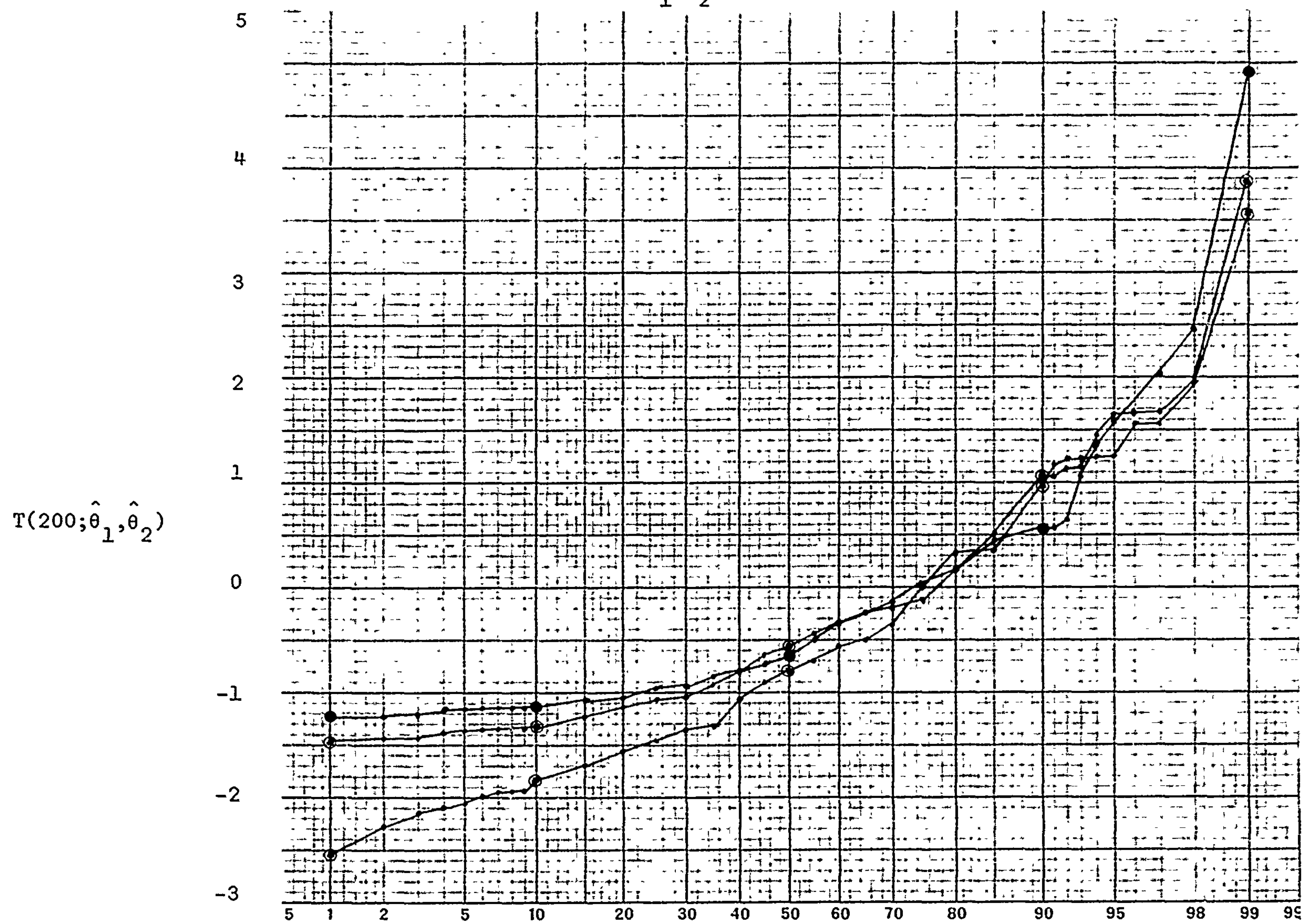


FIGURE A.5
NORMAL PLOT OF ORDERED $T(300;\theta_1,\theta_2)$ VALUES FOR CASES I(⊛),II(◎),III(●)

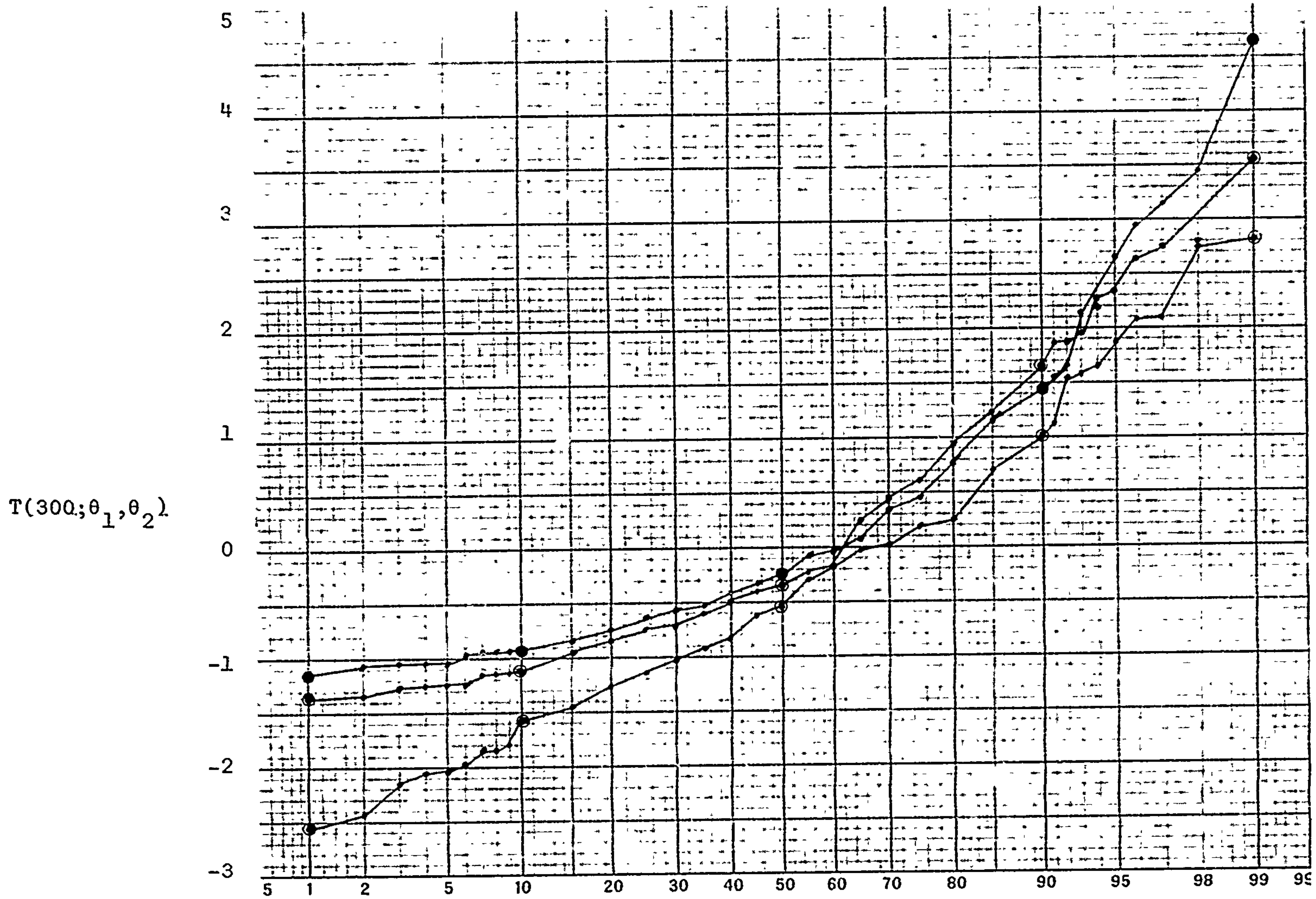
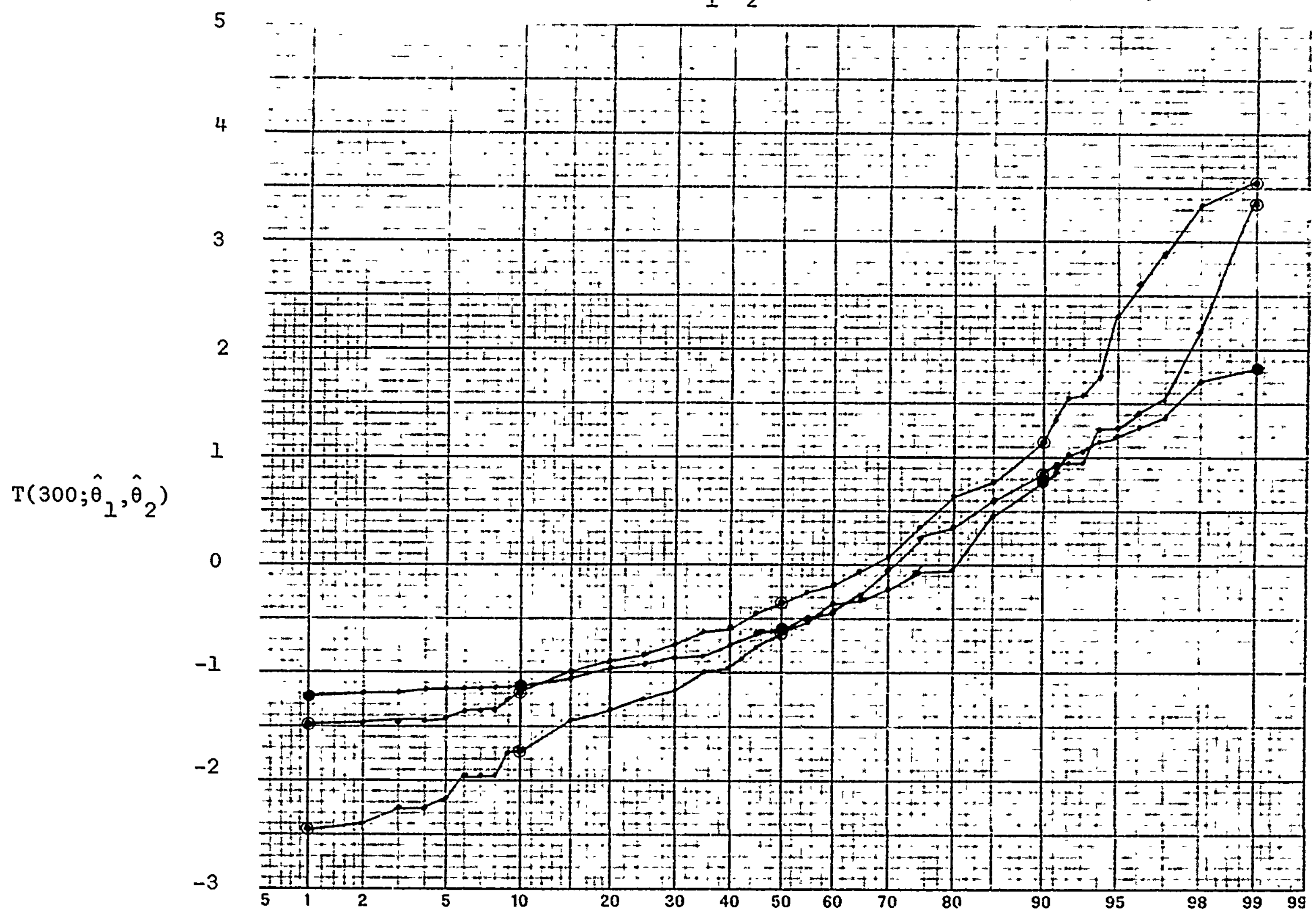


FIGURE A.6
NORMAL PLOT OF ORDERED $T(300;\hat{\theta}_1,\hat{\theta}_2)$ VALUES FOR CASES I(⊛),II(◎),III(●)



APPENDIX B

PROOF OF THEOREM 3.2

Throughout this appendix, we assume that $Y_1(n),\ldots,Y_n(n)$ are the order statistics of a sample from a standard normal population. Suppressing the dependence on n, we let $\tilde{Y}$ denote the (K_n+1)-vector with elements $\tilde{Y}_i = Y_{np_n+(i-1)L_n}$, and let Z denote the (K_n+1)-vector with elements $\tilde{Z}_i = \sqrt{n}t_i(n)(\tilde{Y}_i - T_i(n))$, where n, p_n, L_n, $t_i(n)$, and $T_i(n)$ are as defined in section 3.2.1. The distribution of $\tilde{Z}$ is asymptotically indistinguishable from the (K_n+1)-variate normal distribution with mean vector 0 and covariance matrix $\sum(n)$ is given by (2.23).

Before proving the theorem, it is necessary to establish a linear transformation which can be used to express $\tilde{Z}_i$ as the weighted sum of K_n+1 i.i.d. standard normal random variables.

B.1 A Linear Transformation

Define $U = (U_1,U_2,\ldots,U_{K_n+1})$, where U_i, $1 \leqq i \leqq K_n+1$, are i.i.d. $N(0,1)$ random variables, and let A denote the $(K_n+1)\mathrm{x}(K_n+1)$ matrix with elements

$$a_{ij} = \begin{cases} \sqrt{\ell} & i=j=1 \\ 1 & i=j=2,3\ldots,K_n \\ 1+r & i=j=K_n+1 \\ -1 & (i-1)=j=2,3,\ldots,K_n \\ r\ell^{-1/2} & i=1;j=K_n+1 \\ r & i=2,3,\ldots,K_n;j=K_n+1 \\ 0 & \text{otherwise} \end{cases}$$

where

$$\ell \equiv \frac{L_n}{np_n} \quad \text{and} \quad r \equiv -\frac{L_n(1+p_n^{-1/2})}{n(1-p_n)} \quad .$$

It is easily verified that $A'A = \sum_1^{-1}(n)$ (where $\sum_1^{-1}(n) = (n(L_n-1)/L_n^2)\sum(n)$ is given explicitly in (2.27)). Thus, letting $B = (n(L_n-1)/L_n^2)^{-1/2}A^{-1}$, the random vector $Z^* = BU$ has the same asymptotic distribution as $\tilde{Z}$ and can be used in place of $\tilde{Z}$ for all asymptotic probability calculations.

Let $D_1(m)$ denote the determinant of the $m \times m$ matrix

$$\begin{bmatrix} 1 & & & & r \\ -1 & 1 & & & r \\ & & \ddots & & \vdots \\ & & & 1 & r \\ & & & -1 & 1+r \end{bmatrix}$$

where all remaining elements are zeros. Expanding by the first row,

$$D_1(m) = D_1(m-1) + r(-1)^{m-1}(-1)^{m-1} = D_1(m-1) + r$$
$$= D_1(m-2) + 2r = \ldots = D_1(1) + (m-1)r = 1 + mr .$$

Now call $D(A)$ the determinant of A. Then, expanding by the first row of A,

$$D(A) = \sqrt{\ell}D_1(K_n) + \frac{r}{\sqrt{\ell}}(-1)^{K_n}(-1)^{K_n} = \sqrt{\ell}(1+K_n r) + \frac{r}{\sqrt{\ell}} \tag{B.1}$$

It is now easily verified that the matrix B has elements b_{ij} given by

$$b_{ij} = \left[\frac{L_n^2}{n(L_n-1)}\right]^{1/2} \frac{1}{D(A)} \begin{cases} 1+(K_n+1-i)r & j=1 \\ \sqrt{\ell}(1+(K_n+1-i)r) & i \geq j \geq 1 \\ -r((i-1)\sqrt{\ell} + \frac{1}{\sqrt{\ell}}) & i < j \end{cases} . \tag{B.2}$$

We note that in the "equally spaced" case, $\ell \equiv 1$, and the transformation A reduces to the transformation given in Weiss [38].

B.2 The De Wet and Venter Theorem

The proof of Theorem 3.2 is based on the work of De Wet and Venter [9 ,

10]. For completeness, their key theorem is included here. The notation used in this subsection is their notation and does not refer to any quantities defined elsewhere in this thesis.

Let $Z_1, Z_2, \ldots$ be i.i.d. random variables with zero means, unit variances, and finite fourth moments. Let c_{ijn} $(i,j=1,2,\ldots,n)$ be real numbers with $c_{ijn} = c_{jin}$ for all i and j and set

$$T_n = \sum_{i,j} c_{ijn} Z_i Z_j \quad .$$

Theorem (De Wet and Venter [10]). Let $\{b_{imn}; i=1,2,\ldots,n;\ m=1,2,\ldots\}$ and $\{\gamma_m; m=1,2,\ldots\}$ be real numbers; let $\{M_n\}$ be a sequence of integers with $M_n \to \infty$; let

$$t_n = \sum_{m=1}^{M_n} |\gamma_m| \quad .$$

Assume that as $n \to \infty$,

$$\max_{1<i<n} |b_{imn}| \to 0 \qquad \text{for each } m$$

$$\max_{1\leq k,m\leq M_n} \left|\sum_i b_{imn} b_{ikn} - \delta_{mk}\right| = o(t_n^{-1})$$

$$\max_{1<k,m<M_n} \sum_i b_{imn}^2 b_{ikn}^2 = o(t_n^{-2})$$

$$\sum_{i,j} c_{ijn}^2 \to \Gamma = \sum_1^\infty \gamma_m^2 < \infty$$

$$\max_{1\leq k,m\leq M_n} \left| \sum_{i,j} c_{ijn} b_{imn} b_{jmn} - \gamma_m \right| = o(t_n^{-1}).$$

Then,

$$\text{Dist}(T_n - ET_n) \to \text{Dist}\Big(\sum_1^\infty \gamma_m (Y_m^2 - 1) \Big)$$

with $Y_1, Y_2, \ldots$ i.i.d. $N(0,1)$.

The assumptions of the theorem appear quite formidable to verify. However, as De Wet and Venter point out, it is often the case that

$$c_{ijn} \cong n^{-1}c(i/n+1, j/n+1)$$

where $c(\cdot,\cdot)$ is measurable, square-integrable on the unit square, and symmetric in its arguments. Then, letting $\{g_m\}$ and $\{\gamma_m\}$ denote the eigenfunctions and eigenvalues of c and supposing $\{g_m\}$ is an orthonormal system on $(0,1)$, selecting $b_{imn} = n^{-1/2}g_m(i/n+1)$ will often satisfy the assumptions. Verification involves showing that the differences between the various sums and the corresponding integrals are sufficiently small.

B.3 Proof of Theorem 3.2

We consider the Taylor series expansion of $\tilde{W}^*(n)$ analogous to the expansion (3.19). It is easily seen that the constant term in the new expansion is identically equal to 1; the linear term is identically equal to zero, and the remaining terms are asymptotically exactly the same as in the expansion (3.19) (with $\tilde{T}_i(n)$ replaced by $T_i(n)$ since $q_n \equiv 1$). Thus, as proved in lemma 3.7, the cubic and higher order terms (i.e. $S_4(n)$) are of smaller order in probability than the quadratic term as $n \to \infty$, and the asymptotic distribution of $(1 - W^*(n))$ is thus the same as the asymptotic distribution of $(-S_2(n))$ given by

$$-\frac{L_n}{n}\sum_{i=1}^{K_n+1}\sum_{j=1}^{K_n+1}\left[\frac{L_n}{n}T_i(n)T_j(n) + (K_n+1)^{-1} - \delta_{ij}\right][\tilde{Y}_i - T_i(n)][\tilde{Y}_j - T_j(n)]. \quad (B.3)$$

We define the functions $\{\xi_{ij};\ i,j=1,2,\ldots,K_{n+1}\}$ by

$$\xi_{ij} = \begin{cases} 1 - (p_n + (i-1)\frac{L_n}{n}) & i \geqq j \\ -(p_n + (i-1)\frac{L_n}{n}) & i < j \end{cases}$$

Using the definitions together with the transformation (B.2), tedious but straightforward calculations yield the result that

$$-nS_2(n) \;=\; \frac{L_n}{n}\sum_{i=1}^{K_n+1}\sum_{j=1}^{K_n+1} c_{ij}(n)U_iU_j + R(n) \tag{B.4}$$

where

$$c_{ij}(n) \;=\; c_{ij}^{(1)}(n) - c_{ij}^{(2)}(n) - c_{ij}^{(3)}(n)$$

with

$$c_{ij}^{(1)}(n) = \frac{L_n}{n}\sum_{k=1}^{K_n+1}\frac{\xi_{ki}\xi_{kj}}{t_k^2(n)}$$

$$c_{ij}^{(2)}(n) = \frac{L_n^2}{n^2}\sum_{k=1}^{K_n+1}\sum_{m=1}^{K_n+1}\frac{\xi_{ki}\xi_{mj}}{t_k(n)t_m(n)}$$

$$c_{ij}^{(3)}(n) = \frac{L_n^2}{n}\sum_{k=1}^{K_n+1}\sum_{m=1}^{K_n+1}\frac{T_k(n)T_m(n)}{t_k(n)t_m(n)}\xi_{ki}\xi_{mj}$$

and $R(n)$ is $o_p(p_n^{1/2}(\log 1/p_n)^2)$ as $n \to \infty$. In writing the coefficients for the above sums, we have used the fact that $D^2(A) = \ell r^2 K_n^2(1+o(1))$ as $n \to \infty$.

The asymptotic expected value of $N(1-\tilde{W}^*(n))$ is obviously a_n^*, which was shown in (3.41) to be $o(\log 1/p_n)$ as $n \to \infty$. Futhermore, the expression

$$\frac{L_n}{n}\sum_{i=1}^{K_n+1}\sum_{j=1}^{K_n+1} c_{ij}(n)U_iU_j$$

in (B.4) above is completely analogous to expression (50) of De Wet and Venter [9] with their ψ_{ik} replaced by ξ_{ki}, except that the $\{U_i\}$ here are normally distributed, whereas the $\{Z_i\}$ in their expression (50) are exponentially distributed with mean zero and variance 1. However, the theorem stated in section B.2 depends only on the independence and certain moment conditions of the variables, all of which are satisfied in either

case. Hence, their derivation of the distribution of their statistic r_n^2 [9,pp. 144-5] and [10,pp. 385-6] immediately implies that $n(1-\tilde{W}^*(n)) - a_n^*$ is asymptotically distributed as

$$\sum_{k=3}^{\infty} \frac{X_k^2 - 1}{k}$$

where $X_3, X_4, \ldots$ are i.i.d. standard normal random variables. This completes the proof of theorem 3.2.

In section 3.5 we introduced the statistic $\tilde{\tilde{W}}^*(n)$. Extension of the above proof to the distribution of this statistic requires further elaboration of the method of De Wet and Venter. Set

$$c(x,y) = c^{(1)}(x,y) - c^{(2)}(x,y) - c^{(3)}(x,y)$$

with

$$c^{(1)}(x,y) = \int_0^1 \frac{\xi(z,x)\xi(z,y)}{t^2(z)} dz$$

$$c^{(2)}(x,y) = \int_0^1 \int_0^1 \frac{\xi(z,x)\xi(v,y)}{t(z)t(v)} dxdv = g_1(x)g_2(y)$$

$$c^{(3)}(x,y) = \int_0^1 \int_0^1 \frac{T(z)T(v)}{t(z)t(v)} \xi(z,x)\xi(v,y)dzdv = \frac{1}{2} g_2(x)g_2(y)$$

where

$$\xi(x,y) = \begin{cases} 1-x & \text{if } x \geq y \\ -x & \text{if } x < y \end{cases}$$

and $g_k(x) = H_k(T(x))$ with $H_k(\cdot)$ the k-th normalized Hermite polynomial. The above functions are the continuous analogs of the coefficients in (B.4). The sequences $\{g_1, g_2, \ldots\}$ and $\{1, 1/2, 1/3, \ldots\}$ are the eigenfunctions and eigenvalues of $c^{(1)}(x,y)$, so that $\{g_3, g_4, \ldots\}$ and $\{1/3, 1/4, \ldots\}$ are the

eigenfunctions and eigenvalues of $c(x,y)$. This fact leads to the choices

$$\begin{aligned} b_{imn} &= g_m(p_n+(i-1)L_n/n) \\ \gamma_m &= 1/m \end{aligned} \qquad m = 3,4,\ldots$$

in the De Wet and Venter thoerem, and the result of Theorem 3.2 follows for $\tilde{W}^*(n)$.

For $\tilde{\tilde{W}}^*(n)$, expression (B.3) is unchanged except that the term $(K_n+1)^{-1}$ is missing. This, in turn, implies that (B.4) is unchanged except that $c_{ij}(n) = c_{ij}^{(1)}(n) - c_{ij}^{(3)}(n)$; the terms corresponding to $c_{ij}^{(2)}(n)$ vanish. Thus, in this case, $c(x,y)$ has eigenfunctions $\{g_2,g_3,\ldots\}$ and eigenvalues $\{1/2,1/3,\ldots\}$. It follows that $n(1-\tilde{\tilde{W}}^*(n)) - a_n^{**}$ is asymptotically distributed as

$$\sum_{k=2}^{\infty} \frac{X_k^2 - 1}{k}$$

where $X_2,X_3,\ldots$ are i.i.d. standard normal random variables. This proves part (ii) of corollary 3.3. We named the above distribution $\Psi_1(y)$. Tables of this distribution are not available but can be created by numerically inverting the characteristic function as in [9].

APPENDIX C

AUXILIARY FORMULAS

In this appendix, we collect for reference purposes various formulas concerning functions of $\Phi^{-1}(x)$ which were required in chapters 2 and 3. We also illustrate the formal procedure for approximating summation by integration and obtain upper bounds on the errors of such approximations.

C.1 Derivatives, Integrals and Limits for Functions of $\Phi^{-1}(x)$

Derivatives. Let $T(x)$ denote $\Phi^{-1}(x)$, $t(x)$ denote $\phi(T(x))$, and $f(x)$ denote $t(x)T(x)$. Then the first four derivatives of $f(x)$ are given by

$$\begin{aligned} f^{(1)}(x) &= 1 - T^2(x) \\ f^{(2)}(x) &= -2T(x)/t(x) \\ f^{(3)}(x) &= -2(1 + T^2(x))/t^2(x) \\ f^{(4)}(x) &= -4(T^3(x) + T(x))/t^3(x). \end{aligned} \tag{C.1}$$

It is easily verified by induction that as x approaches 0 or 1, we have

$$f^{(k)}(x) = -2\cdot(k-2)!\,T^{k-1}(x)/t^{k-1}(x) + \text{(smaller order terms)}$$

for $k \geq 2$.

Integrals. The analysis in chapter 3 required evaluation of expressions of the form

$$c(n) \int_{a(n)}^{b(n)} \frac{T^k(x)}{t^k(x)}\, dx \tag{C.2}$$

where $k \geq 2$ is an integer, $a(n) \to 0$ and $b(n) \to 1$ as $n \to \infty$, and where $c(n)$ approaches zero rapidly enough that the expression remains bounded. Using the fact that $\frac{d}{dx} t^{-1}(x) = T(x)/t^2(x)$, we can obtain an iterative relation for integrals of the form (C.2) by integrating by parts. We have

$$\int_{a(n)}^{b(n)} \frac{T^k(x)}{t^k(x)}\, dx = \int_{a(n)}^{b(n)} \frac{T^{k-1}(x)}{t^{k-2}(x)} \frac{T(x)}{t^2(x)}\, dx$$

$$= \frac{T^{k-1}(x)}{t^{k-1}(x)} \Big|_{a(n)}^{b(n)} - \int_{a(n)}^{b(n)} \{(k-1) \frac{T^{k-2}(x)}{t^k(x)} + (k-2) \frac{T^k(x)}{t^k(x)}\} dx,$$

so that

$$\int_{a(n)}^{b(n)} \frac{T^k(x)}{t^k(x)}\, dx = \frac{1}{k-1} \frac{T^{k-1}(x)}{t^{k-1}(x)} \Big|_{a(n)}^{b(n)} - \int_{a(n)}^{b(n)} \frac{T^{k-2}(x)}{t^k(x)}\, dx. \tag{C.3}$$

Similarly,

$$\int_{a(n)}^{b(n)} \frac{T^{k-2j}(x)}{t^k(x)}\, dx = \frac{1}{k-1} \frac{T^{k-2j-1}(x)}{t^{k-1}(x)} \Big|_{a(n)}^{b(n)} - \frac{(k-2j-1)}{k-1} \int_{a(n)}^{b(n)} \frac{T^{k-2j-1}(x)}{t^k(x)}\, dx$$

for $j = 1,2,\ldots,[k-1/2]$ (where $[h]$ denotes the largest integer $\leq h$), and

$$\int_{a(n)}^{b(n)} \frac{T(x)}{t^k(x)}\, dx = \frac{1}{k-1} \frac{1}{t^{k-1}(x)} \Big|_{a(n)}^{b(n)} . \tag{C.4}$$

Combining the above, the integral in (C.2) can be written

$$\frac{1}{k-1}\, t^{1-k}(x)\{T^{k-1}(x) + \sum_{j=1}^{[\frac{k-1}{2}]} (-1)^j \{\frac{(k+1-2j)(k+3-2j)\cdots(k-1)}{(k-1)^j}\} T^{k-2j-1}(x)\}\Big|_{a(n)}^{b(n)}$$

$$+ \frac{1}{2}(1+(-1)^k)\, \frac{1\cdot 3\cdot 5\cdots(k-1)}{(k-1)^{k/2}} \int_{a(n)}^{b(n)} t^{-k}(x)dx. \tag{C.5}$$

Note that is $a(n) \to 0$ and $b(n) \to 1$ as $n \to \infty$, then the term

$$\frac{1}{k-1}\, \frac{T^{k-1}(x)}{t^{k-1}(x)}\Big|_{a(n)}^{b(n)}$$

dominates the above expressions. Also note that if k is odd and $a(n) = 1-b(n)$, then (C.2) is zero by symmetry. Finally, from (C.4),

$$\begin{aligned}\int_{a(n)}^{b(n)} t^{-2}(x)dx &= o(t^{-1}(a(n)) + t^{-1}(b(n))) \\ &= o([a^2(n)\log 1/a(n)]^{-1/2} + [(1-b(n))^2 \log 1/(1-b(n))]^{-1/2})\end{aligned} \tag{C.6}$$

if $a(n) \to 0$ and $b(n) \to 1$ as $n \to \infty$, where the last line follows from (C.10) below.

<u>Limits</u>. Using the fact [13,p. 175] that

$$1 - \Phi(x) \sim \frac{\phi(x)}{x} \quad \text{as} \quad x \to \infty$$

and taking $x = \Phi^{-1}(y)$ (and $y \to 1$), we obtain

$$\frac{T(y)}{t(y)} \sim \frac{1}{1-y} \quad \text{as} \quad y \to 1 \ . \tag{C.7}$$

The following are also easily verified:

$$\lim_{y\to 0} \frac{T^2(y)}{\log(1/y)} = 2 \tag{C.8}$$

$$\lim_{y\to 0} \frac{t(y)T(y)}{y\log(1/y)} = -2 \tag{C.9}$$

$$\lim_{y\to 0} \frac{t(y)}{y(\log 1/y)^{1/2}} = \sqrt{2} \ . \tag{C.10}$$

The corresponding limits as $y \to 1$ are obvious.

C.2 Approximating Summation by Integration

For concreteness, we formalize the procedure used throughout chapters 2 and 3 using two expressions which are representative of those appearing in chapter 3.

(A) Let $\psi(x)$ denote the function $T^j(x)/t^k(x)$, where k is a positive integer and j is an even integer $\geq k$. Also let a_{ni} denote the quantity $p_n + (i-1)L_n/n$, and set $i^* =$ greatest integer $\leq K_n/2$. Let $S(n)$ denote the sum

$$\frac{L_n}{n} \sum_{i=2}^{K_n} \psi(a_{ni})$$

and let $I(n)$ denote the integral

$$\int_{p_n}^{1-p_n} \psi(x)dx .$$

Then, $S(n) - I(n) = E_1(n) + E_2(n) + E_3(n)$ where

$$E_1(n) = \frac{L_n}{n} \sum_{i=2}^{i*} \psi(a_{ni}) - \int_{p_n}^{a_{ni*}} \psi(x)dx$$

$$E_2(n) = \frac{L_n}{n} \sum_{i=i*+1}^{K_n} \psi(a_{ni}) - \int_{a_{n(i*+1)}}^{1-p_n} \psi(x)dx$$

$$E_3(n) = - \int_{a_{ni*}}^{a_{n(i*+1)}} \psi(x)dx .$$

Using the fact that $\psi(x)$ is monotonically decreasing on $(0,1/2)$,

$$|E_1(n)| \leqq \sum_{i=2}^{i*} \int_{a_{n(i-1)}}^{a_{ni}} |\psi(a_{ni}) - \psi(x)|dx$$

$$\leqq \sum_{i=2}^{i*} \int_{a_{n(i-1)}}^{a_{ni}} |\psi(a_{ni}) - \psi(a_{n(i-1)})|dx$$

$$= \frac{L_n}{n} \sum_{i=2}^{i*} (\psi(a_{n(i-1)}) - \psi(a_{ni}))$$

$$= \frac{L_n}{n} (\psi(a_{n1}) - \psi(a_{ni*}))$$

$$= O(\frac{L_n}{n})\ p_n^{-k}(\log 1/p_n)^{(j-k)/2})\ .$$

A symmetric argument applies to $|E_2(n)|$. Clearly, $E_3(n)$ is $O(L_n/n)$. Thus $|S(n) - I(n)|$ is $O((L_n/n)p_n^{-k}(\log 1/p_n)^{(j-k)/2})$. From (C.5), $I(n)$ is $O(p_n^{1-k}(\log 1/p_n)^{(j-k)/2})$. Hence,

$$S(n) = I(n)(1 + O(L_n/np_n)).$$

(B) Let $\bar{S}(n)$ denote

$$\frac{L_n}{n} \sum_{i=2}^{K_n} [\frac{\varepsilon_n(T(a_{ni}))}{t(a_{ni})}]^2$$

and let $\bar{I}(n)$ denote

$$\int_{p_n}^{1-p_n} [\frac{\varepsilon_n(T(x))}{t(x)}]^2 dx\ .$$

Then,

$$|\bar{S}(n) - \bar{I}(n)| \leq \sum_{i=2}^{K_n} \int_{a_{n(i-1)}}^{a_{ni}} |[\frac{\varepsilon_n(T(a_{ni}))}{t(a_{ni})}]^2 - [\frac{\varepsilon_n(T(x))}{t(x)}]^2 |dx\ .$$

Since

$$\left[\frac{\varepsilon_n(T(x+(L_n/n)))}{t(x+(L_n/n))}\right]^2 = \left[\frac{\varepsilon_n(T(x))}{t(x)}\right]^2 + 2\,\frac{L_n}{n}\,\frac{\varepsilon_n(T(\xi))}{t(\xi)}\left[\frac{\varepsilon_n'(T(\xi))}{t^2(\xi)} + \frac{\varepsilon_n(T(\xi))}{t(\xi)}\,\frac{T(\xi)}{t(\xi)}\right]$$

for some ξ, $x < \xi < x+(L_n/n)$, (3.62) and (3.63) imply that

$$\begin{aligned}|\bar{S}(n) - \bar{I}(n)| &= o\left(2\,\frac{L_n}{nb^2(n)}\,(\log 1/p_n) + 2\,\frac{L_n}{nb^2(n)}\sum_{i=2}^{K_n}\left|\frac{T(a_{ni})}{t(a_{ni})}\right|\right)\\ &= o\left(\frac{L_n}{nb^2(n)}\,(\log 1/p_n)\right)\end{aligned}$$

where $b(n)$ is as defined in (3.62) and (3.63). In particular, the latter expression is $o(\rho(n,\varepsilon_n))$. Similar results may be obtained for the integrals of Chapter 2 using (2.41).

BIBLIOGRAPHY

[1] Bartle, R.G. (1964), The Elements of Real Analysis, John Wiley and Sons, New York.

[2] Blumenthal, S. (1966), "Contributions to Sample Spacings Theory, II; Tests of the Parametric Goodness of Fit and Two-Sample Problems", Ann. Math. Stat., Vol. 37, pp. 925-939.

[3] Chen, E.H. (1971), "The Power of the Shapiro-Wilk W-Test for Normality in Samples from Contaminated Normal Distributions", Jour. Amer. Stat. Assoc., Vol. 66, pp. 760-762.

[4] Cibisov, D.M. (1961), "On the Tests of Fit Based on Sample Spacings", Theoria vero. i. Prim., Vol. 6, pp. 354-358, Theory Probab. Applic., Vol. 6, pp. 325-329.

[5] D'Agostino, R.B. (1971), "An Omnibus Test of Normality for Moderate and Large Size Samples", Biometrika, Vol. 58, pp. 341-347.

[6] Darling, D.A. (1955), "The Cramér-Smirnov Test in the Parametric Case", Ann. Math. Stat., Vol. 26, pp. 1-20.

[7] Darling, D.A. (1957), "The Kolmogorov-Smirnov, Cramér-von Mises Tests", Ann. Math. Stat., Vol. 28, pp. 823-838.

[8] David, H.A. (1970), Order Statistics, John Wiley and Sons, New York.

[9] De Wet, T. and Venter, J.H. (1972), "The Asymptotic Distribution of Certain Test Criteria of Normality", S. Afr. Statist. J., Vol. 6, pp. 135-149.

[10] De Wet, T. and Venter, J.H. (1973), "The Asymptotic Distributions for Quadratic Forms with Application to Tests of Fit", Annals of Statistics, Vol. 1, pp. 380-387.

[11] Durbin, J. (1971), "Boundary-Crossing Probabilities for the Brownian Motion and Poisson Processes and Techniques for Computing the Power of the Kolmogorov-Smirnov Test", J. Appl. Probability, Vol. 8, pp. 431-453.

[12] Durbin, J. (1975), "Kolmogorov-Smirnov Tests When Parameters Are Estimated with Applications to Tests of Exponentiality and Tests on Spacings", Biometrika, Vol. 62, pp. 5-22.

[13] Feller, W. (1968), An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Edition, John Wiley and Sons, New York,

[14] Celler, M.(1963), "A Table of Integrals Involving Powers, Exponentials, Logarithms, and the Exponential Integral", Technical Report No. 32-469, Jet Propulsion Laboratory.

[15] Kac, M., Kiefer, J. and Wolfowitz, J. (1955), "On Test of Normality and Other Tests of Goodness of Fit Based on Distance Methods", Ann. Math. Stat., Vol. 26, pp. 189-211.

[16] Lilliefors, H. W. (1967), "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown", Jour. Amer. Stat. Assoc., Vol. 62, pp. 399-402.

[17] Lilliefors, H.W. (1969), "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown", Jour. Amer. Stat. Assoc., Vol. 64, pp. 387-389.

[18] Massey, F.J. Jr. (1950), "A Note on the Power of a Non-parametric Test", Ann. Math. Stat., Vol. 21, pp. 440-443.

[19] Moore, D.S. and Spruill, M.C. (1975), "Unified Large-Sample Theory of General Chi-Squared Statistics for Tests of Fit", Annals of Statistics, Vol. 3, pp. 599-616.

[20] Pyke, R. (1965), "Spacings", J. R. Stat. Soc. B, Vol. 3, pp. 395-449.

[21] Reiss, R.D. (1974), "The Asymptotic Normality and Asymptotic Expansions for the Joint Distribution of Several Order Statistics", Preprints in Statistics No. 11, University of Cologne.

[22] Sarkadi, K. (1975), "The Consistency of the Shapiro-Francia Test", Biometrika, Vol. 62, pp. 445-450.

[23] Sen, P.K. (1959), "On the Moments of the Sample Quantiles", Calcutta Stat. Assoc. Bull., Vol. 9, pp. 1-19.

[24] Serfling, R.J. and Wood, C.L. (1975), "On Null-Hypothesis Limiting Distributions of Kolmogorov-Smirnov Type Statistics with Estimated Location and Scale Parameters", Florida State University Statistics, Report M339, Department of Statistics, Florida State University.

[25] Sethuraman, J. and Rao, J.S. (1970), "Pitman Efficiencies of Tests Based on Spacings", in Nonparametric Techniques in Statistical Inference, M.L. Puri (ed.), Cambridge University Press, London, pp. 405-415.

[26] Shapiro, S.S. and Francia, R.S. (1972), "An Approximate Analysis of Variance Test for Normality", Jour. Amer. Stat. Assoc., Vol. 67, pp. 215-216.

[27] Shapiro, S.S. and Wilk, M.B. (1965), "An Analysis of Variance Test for Normality (Complete Samples)", Biometrika, Vol. 52, pp. 591-611.

[28] Shapiro, S.S. and Wilk, M.B. (1968), "Approximations for the Null Distribution of the W Statistic", Tech., Vol. 10, pp. 861-866.

[29] Shapiro, S.S., Wilk, M.B., and Chen, H.J. (1968), "A Comparative Study of Various Tests for Normality", Jour. Amer. Stat Assoc., Vol. 63, pp. 1343-1370.

[30] Stephens, M.A. (1974), "EDF Statistics for Goodness of Fit and Some Comparisons", Jour. Amer. Stat. Assoc., Vol. 69, pp. 730-737.

[31] Stephens, M.A. (1975), "Asymptotic Properties for Covariance Matrices of Order Statistics", Biometrika, Vol. 62, pp. 23-28.

[32] Stephens, M.A. (1975), "Extensions to the W-Test for Normality", Technical Report No. 215, Department of Statistics, Stanford University.

[33] Stephens, M.A. (1975), "On the Distributions of the W and W' Statistics for Testing for Normality", Technical Report No. 218, Department of Statistics, Stanford University.

[34] Stephens, M.A. (1976), "Asymptotic Results for Goodness-of-Fit Statistics with Unknown Parameters", Annals of Statistics, Vol. 4, pp. 357-369.

[35] Theune, J.A. (1973), "Comparison of Power for the D'Agostino and the Wilk-Shapiro Test of Normality for Small and Moderate Samples", Stat. Neerlandica, Vol. 27, pp. 163-169.

[36] Thompson, R. (1966), "Bias of the One-Sample Cramér-von Mises Test", Jour. Amer. Stat. Assoc., Vol. 61, pp. 246-247.

[37] Weiss, L. (1962), Review of Cibisov (1961), Math. Rev., Vol. 24, No. A1776.

[38] Weiss, L. (1969), "The Asymptotic Joint Distribution of an Increasing Number of Sample Quantiles", Ann. Inst. Stat. Math., Vol. 21, pp. 257-263.

[39] Weiss, L. (1973), "Statistical Procedures Based on a Gradually Increasing Number of Order Statistics", Communications in Statistics, Vol. 2 (2), pp. 95-114.

[40] Weiss, L. (1974), "The Asymptotic Sufficiency of a Relatively Small Number of Order Statistics in Tests of Fit", Annals of Statistics, Vol. 2, pp. 795-802.

[41] Weiss, L. (1975), "Testing Fit with Nuisance Location and Scale Parameters", Naval Research Logistics Quarterly, Vol. 22, pp. 55-63

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
#310

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
GOODNESS OF FIT TESTS FOR COMPOSITE HYPOTHESES BASED ON AN INCREASING NUMBER OF ORDER STATISTICS

5. TYPE OF REPORT & PERIOD COVERED
Technical Report

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Ray Hamilton Jakobovits

8. CONTRACT OR GRANT NUMBER(s)
DAHC04-73-C-0008
N00014-75-C-0586 Add?
NSF GRANT NO. MCS76-06340

9. PERFORMING ORGANIZATION NAME AND ADDRESS
School of Operations Research & Industrial Engineering, College of Engineering
Cornell University, Ithaca NY 14853

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Sponsoring Military Activity
U.S. Army Research Office, P.O. Box 12211
Durham, N.C. 27706

11. CONTROLLING OFFICE NAME AND ADDRESS
Sponsoring Military Activity
Statistics and Probability Program
Office of Naval Research
Arlington, VA. 22217

12. REPORT DATE
September 1976

13. NUMBER OF PAGES
158

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Goodness of Fit Tests; Tests for Normality; Asymptotic Distribution Theory; Order Statistics.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

~~We are~~ is shown concerned with sequences of tests ~~of the composite~~ null hypothesis that the distribution function of an observed sample is a member of a specified scale-location parameter family, where the scale and location parameters are unknown and unspecified. ~~We~~ first, ~~study~~ a test of this hypothesis proposed by L. Weiss is studied. Then ~~we propose and analyze~~ several tests for the most common special case--that of normality are analyzed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

All of the results are asymptotic, as the sample size n tends to infinity, and all of the test criteria which we examine are based on functions of an increasing subset of the ordered sample values. By increasing subset, we mean that as the sample size tends to infinity, the number of order statistics which we consider also tends to infinity. By assuming certain relationships among the number of sample quantiles in the selected subset, the separation between those quantiles, the rate at which the quantiles move into the tails of the distribution, and the smoothness and size of the tails of the density function of a standard representative of the hypothesized scale-location parameter family, we are able to assert that the selected subset of sample quantiles is asymptotically jointly normally distributed with known mean vector and covariance matrix.

The Weiss procedure is based on a quadratic form of the sample quantiles which has an asymptotic chi-square distribution. We examine the asymptotic power of his test under sequences of contiguous alternatives of the form $H_n(x) = G((x-\theta_1)/\theta_2) + \varepsilon_n(x)$, where $G(y)$ is the standard representative of the scale-location parameter family appearing in the null hypothesis, and the "disturbance functions" $\{\varepsilon_n(x)\}$ satisfy certain regularity conditions. By letting the sequence $\{\varepsilon_n(x)\}$ approach zero at a certain rate and with respect to a certain measure of distance, we obtain a non-trivial power for the test.

Turning to the composite hypothesis of normality and using the same approach of selecting a gradually increasing subset of order statistics, we develop large-sample analogs of the Wilk-Shapiro and Shapiro-Francia tests, two of the most effective (small samples) but also least understood tests for normality. We show that the analog of the Wilk-Shapiro statistic is asymptotically normally distributed under the null hypothesis as the sample size tends to infinity. We also show that, asymptotically, the analog of the Shapiro-Francia statistic, suitably standardized, has the same distribution under the null hypothesis as a certain weighted sum of independent chi-square random variables. Up to a shift in location, the latter distribution is the same as that derived by de Wet and Venter for an analog of the Shapiro-Francia test based on the complete set of order statistics. We further prove that the tests based on the above statistics are consistent. Finally, we study the behavior of those tests under the sequence of alternatives $\{H_n(x)\}$. We find that the measures of distance for the tests are quite complicated. We also find that, contrary to most small-sample empirical studies, the analog of the Shapiro-Francia test has better asymptotic power than the analog of the Wilk-Shapiro test, and we show by example that the analog of the Wilk-Shapiro test can be biased, even against sequences of contiguous alternatives. As a consequence of the power studies, we propose and analyze two further statistics, modifications of those discussed above, which yield improved tests for normality.

DOCUMENT TRANSMITTAL AND RECEIPT
NAVEXOS-2435 (7-67)

DEPARTMENT OF THE NAVY
OFFICE OF NAVAL RESEARCH
WASHINGTON, D. C. 20360

TO

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

MAILING RECORD

DATE 8/31/76

REGISTER NO.

SERIAL NO.

REFERENCE (Identify source and date of request, contract number or related work project)

DOD Instruction 5100.38

The material listed below is forwarded for the purpose checked. Please acknowledge receipt.

☐ INFORMATION AND RETURN BY (date or event) ______ ☒ RETENTION* ☐ RETURN OF LOAN ☐ OTHER (See "Remarks")

QTY	DESCRIPTION OF ENCLOSURES	COPY NO.	SECURITY CLASS.
12	GOODNESS OF FIT TESTS FOR COMPOSITE HYPOTHESES BASED ON AN INCREASING NUMBER OF ORDER STATISTICS Technical Report No. 310, September 1976. Ray Hamilton Jakobovits.		UNCLASSIFIED
	Sch. of Operations Research & Industrial Engineering Upson Hall, Cornell University Ithaca, New York 14853		
	Robert Bechhofer Robert E. Bechhofer, Project Director		

REMARKS

Please Advise:
(1) AD Number Assigned ______
(2) Division Number Assigned ______

SECURITY CLASSIFICATION OF THIS FORM

WITH ENCLOSURES	WITHOUT ENCLOSURES
UNCLASSIFIED	UNCLASSIFIED

TRANSMITTAL AUTHORIZED BY (Signature over typed name)

*NOTE: Upon completion of contract(s) or when material has served its purpose, destroy in accordance with paragraph 14 of Industrial Security Manual for safeguarding classified information.

RETURN TO

OFFICE OF NAVAL RESEARCH (CODE 436)
DEPARTMENT OF THE NAVY
17TH AND CONSTITUTION AVENUE, N. W.
WASHINGTON, D. C. 20360

Receipt of the above listed enclosures is acknowledged.

(Signature and date)

TITLE OF RECIPIENT

NAME OF RECEIVING ACTIVITY

COPY TO: